

No next stop for NATO

This week's meeting of the Council of NATO makes it necessary that there should be a full-throated discussion of what happens next, which means problems for everybody.

NOT so long ago, at the end of 1987, it seemed that Europe had become a place free from nuclear weapons. Now, as this week's meeting of the Council of the North Atlantic Treaty Organisation (NATO) will have shown, Europe seems once more to have become a nuclear armed camp, or a place with ambitions in that direction. What has gone wrong, or otherwise happened, since last November's summit meeting between the United States and the Soviet Union in Washington? Several confusing and even conflicting tendencies are under way, and are likely to continue for several years. The general interest requires that they should be understood as fully as may be possible.

The most obvious of the changing ingredients of the European problem is the arms control process, now mercifully resumed. Last November's treaty (INF) requiring the removal from Europe of missiles of intermediate range is the first to have emerged from five years of negotiation at Geneva because, for technical reasons, it is the simplest that could have been reached. Since the 1950s, when missiles were first developed to carry nuclear warheads, they have become almost conspicuous. By contrast, it is much more difficult to tell what uses are planned for other kinds of delivery systems, aircraft or battle-field howitzers, for example. It is true that President Reagan and Mr Gorbachev might just as well have begun with an agreement on strictly strategic weapons, on which their representatives at Geneva are even now negotiating, except that the strategic issue is unavoidably enmeshed with the US devotion to the Strategic Defense Initiative (SDI).

On that reading, the INF agreement may have been a way of patching up the previous year's failure at Reykjavik, a way of showing that arms control was not dead and (crucial from Mr Gorbachev's point of view) a way of discovering how the US Congress now regards bilateral arms control agreements. Reagan and Gorbachev may have been willing to overlook a few gaps in the document they signed, such as the ambiguous status of British and French nuclear weapons, in the expectation that more agreements would soon follow. Where they appear to have miscalculated is their belief that others would be prepared to put their anxieties on ice until there is another draft treaty from Geneva.

In the event, as this week's meeting will have shown, there are two quite distinct tendencies within Europe proper. West Germany, filled with anxiety at the prospect of INF in advance of last November's summit, but having seen Soviet missiles being dismantled and carted away from Eastern Europe in the past few weeks, does not now relish the notion that short-range nuclear weapons (not covered by INF) might be modernized. (Those who guess that West Germany's motives centre on fears that West Germany would be the likely target for these weapons reckon without the glittering attraction of *Ostpolitik*.) At the same time, Britain and France, closer together on this issue than at any time since 1954, appear determined to keep their independent nuclear forces, in reality substitutes for some of the missiles being dismantled under INF, strong — if not to make them stronger. Finally, the complex political animal which is the United States, having agreed to withdraw nuclear missiles from Europe, is plainly wondering whether it should continue to keep

300,000 members of its armed services there.

All this is understandable, and should also have been predictable. The only course that NATO can reasonably hope to follow is to keep to something like the *status quo* at least until the next instalment of the summit meetings, possibly in April. At that stage, if the planned agreement on strategic arms is to materialize, it is inevitable that British and French nuclear weapons will somehow have to be counted (as, implicitly, they were in the SALT II agreement of 1978). The snag, for France and, increasingly, for Britain is that these independent nuclear forces are justified by the Gallois doctrine of the 1950s that the destructive power of a smaller power's nuclear forces will suffice as a deterrent to greater powers if they match its own value to an adversary as a prize. Although the British Prime Minister, Mrs Margaret Thatcher, has said that British nuclear forces could be negotiable if there were general progress towards arms control (the French have said nothing), it is inevitable that the British nuclear forces will seem more endearing to those who pay for them as those circumstances come closer.

That is why the best strategy, for Britain and France in one camp, West Germany in another but also for the United States, conscious for the first time that it is overstretched by its commitments, is to plan for yet another series of arms control negotiations with the Soviet Union, dealing with European security. That is the only setting in which understanding could rationally be reached on the deployment of short-range nuclear weapons as well as on the British and French strategic nuclear forces. Both sides have much to gain from such an accommodation, which would at least be easier to define than will be the search for a trade-off between strategic arms and SDI. The stumbling block is the old-fashioned truth that an agreement on European security will require an understanding on relations between Eastern European states and the rest of Europe. That is evidently a more serious obstacle for the Soviet Union than for, say, West Germany. □

Electric company sale

The British government has a bad recipe for selling a public monopoly.

THE British government has at least one thing to its credit: it has already explored more ways of selling public monopolies to the private sector than any other government. The now familiar problem is that of balancing the proceeds (which help to swell the Treasury) against the public interest (which requires that newly privatized companies should efficiently sell their services cheaply). Evidently the hope is that some future formula will prove to be a huge success, capable of being replicated indefinitely. But what if the circle cannot be squared, and if there is no basis on which public monopolies made private can be structured in such a way that they are inherently efficient and cheap? In such a case, the British will run out of nationalized industries before the search for the ideal formula is complete.

That is the worry that should be keeping Mr Cecil Parkinson, now the Secretary of State for Energy in the British government,



awake at nights. With remarkable speed, Parkinson last week published a white paper specifying the government's plans for selling off the publicly owned electricity industry, thereby silencing the people now running the industry who believe the sale should be managed differently (see *Nature* **331**, 466; 1988). But even if Parkinson has won the battle, he may not have won the war. The weakness of the white paper is that too many of its promises may be kept only with the greatest difficulty.

What is now proposed is that the present electricity distribution organizations (called "area boards") should become private companies and that they should between them own the British national grid, which effectively sews all present consumers and all present generating stations into a single network. The grid has been constructed and is now operated by the Central Electricity Generating Board (CEGB), which also operates the nationalized generating plant. Meanwhile, the CEGB is to be broken for the purposes of the sale into two parts, the larger of which will own Britain's nuclear power stations and whose business will be to some extent underpinned by the requirement that the distribution companies should buy at least a proportion (not yet fixed) of their electricity from nuclear plants. The distribution companies will be able to buy electricity from other sources, such as novel power generation companies, either directly or through the grid, and will also be able to generate their own electricity if they choose. One small but important innovation is the requirement that the new owners of the grid should let it be used by outsiders, a large customer wishing to buy electricity from other sources, for example.

The plan is radical enough to satisfy the harshest critics, but is it wise? It would be easier to tell if there were some evidence in the white paper that a small army of economists had been asked to anticipate the problems that will arise. Relationships between distribution companies and power generators are to be regulated by contracts freely arrived at, but distribution companies will have an obligation to supply their customers (the other side of the monopoly coin). What will happen if a distribution company reckons that the only safe way of meeting the obligation is to build enough generating capacity to meet its own needs, turning its back on the national grid? In the long run, that is the way in which the present public monopoly will be converted into a dozen private monopolies. Yet how could such a company be justly penalized? It will be virtually impossible to devise a set of regulations that can be declared in advance and are more sophisticated than the familiar yardsticks of return on capital, and which are devices for giving companies a steady source of income.

By far the better scheme would have been to leave the grid with CEGB, but to insist on its use as a common carrier and to allow competition in the construction and operation of generating plant, either for local or national supply. But the government, having nailed its colours to the mast, will not now pull them more than half way down. That is why the best solution now is to separate the grid from both the distribution companies and CEGB, making it into a privately owned power broker, with an interest in making an efficient market in electricity. Meanwhile, CEGB could be made into a company owning all existing nationalized power stations, but with the understanding that decisions about which generating plant next to build would be left to those willing to risk their capital — either CEGB or an upstart competitor. That way, it would not even be necessary to decree a quota for nuclear power, which is hardly the way in which a government "determined that public confidence in the nuclear programme should be maintained" should seek to protect the nuclear industry. Why not leave that to the market for shares in the new companies, and to the profitability of CEGB's successors? For the plain truth is that a single successor to CEGB without a guaranteed sale of electricity from uranium but with potential competitors all around would have to look to efficiency for survival. If it had to build nuclear power stations in the process, that would be its own affair. □

Too many places?

The next reorganization of British academic life will be serious; there should be a simpler way.

THE University Grants Committee (UGC) will be telling British universities, later in the week, where it thinks they stand in what are called the earth sciences, that amalgam of geology with geophysics together with what used to be simple oceanography and/or atmospheric science. This is the culmination of a process under way for two years, punctuated by the appearance of a discussion document by Professor Ron Oxburgh a year ago, and which broke new ground in the management of the British university system by advocating open recognition that some university departments are first class, that others teach well and that still others have no continuing claim on public, let alone academic, attention. Since then, UGC has backtracked a little in the face of the simple logic that it cannot formally instruct universities to forswear some field of study, or to insist that they should never become more than merely competent in the field. UGC appears to be hoping that it has done enough quiet talking that even those who are surprised by the contents of the letters in the post to them will not feel aggrieved. Surely, by now, the argument goes, everybody must understand that there cannot be an excellent department of the earth sciences in every university?

The answer is probably "yes", even if people follow by asking "but why pick on us?". Again, UGC seems to be counting for consent on the care with which it plans that academics with something to contribute should be helped to do so by being given access to neighbouring universities' equipment (and free time in which to do so). But it would naturally be different if it were proposed that a particular university in Britain should sink back into a teaching role in, for example, English literature or physics and/or chemistry? Yet just that crisis may soon burst over the heads of British universities. Panels intended to tell sheep from goats are already at work in both the sciences. Ambitions that the physics panel should report in April are cherished mostly by its chairman, but by the beginning of the next academic year on 1 August, there should be a cat among both sets of pigeons. However generous the arrangements for sharing other institutions' equipment, what when academic physicists and chemists are told, well in time for 1990–91, that their bread and butter depends on service teaching, or even nothing much?

The riposte, that it had better not come to that, is also a good touchstone to judge what happens next. However enthralled outsiders may have been by the long agony of British academic life, they may not appreciate that the decisions now being made about the character of particular university departments are being made by academics in what they regard as the best interests of the system to which they belong. The decisions are forced by the money problem to which people have become accustomed, but which is an externality. The Education Reform Bill is formally irrelevant (but may make things even worse).

What, in the circumstances, should happen? The hoped-for recognition among earth scientists that reorganization is inevitable had better be translated to a cruder and simpler language. Britain as it is has too many universities, largely because the Robbins Committee in 1963 advocated a three-fold increase of the number of students coupled with a doubling of the number of institutions. UGC's present administration of that state of affairs (complicated by the equipotent polytechnics) cannot shrink the number of institutions; its best hope is that some of them will find better livings as specialized institutions called universities only in their names. But it would be better if that nettle could be grasped, and quickly. There are no more institutions that need to be sent into oblivion, and hardly any academics that need in the national interest to be fired, but there is a need that a government that pretends to want to run the university system should help it to become what it might be. That might even, in due course, be called education reform. □

AIDS commission report confounds critics

- Emphasis on treatment and research
- Almost 200 recommendations made

Washington

THE US Presidential Commission on the HIV Epidemic surprised observers this week by sending to the White House detailed preliminary recommendations for dealing with the AIDS (acquired immune deficiency syndrome) epidemic that go well beyond the soft approach favoured by the Reagan administration.

The commission's 180 specific recommendations were included in a 60-page report spelling out three priority areas: intravenous drug abuse treatment; more flexible patient care; and higher priorities for basic research and drug development programmes. The commission originally intended to assess the prevalence of HIV infection in the United States, but this will now be deferred to September.

The 13-member commission issued its recommendations earlier this week after three months of hearings and data-gathering. More than 350 witnesses testified, including AIDS patients, health-care practitioners, gay activists, educators, researchers and government agencies.

The keystone of the commission's recommendations for countering intravenous drug abuse is "treatment on demand". There are now roughly 1.2 million intravenous drug users in the United States, but only 148,000 are receiving treatment at any given time. In the 24 cities with the largest numbers of drug abusers, addicts must wait up to six months for treatment. The commission estimates that 3,300 new drug treatment centres will be needed to accommodate these people, with a projected cost of \$1,500 million. Half of the money is expected to come from the federal government, with the balance being made up by state and local governments.

To address the unique needs of those infected with HIV — who may require only intermittent hospitalization for acute opportunistic infections until they become gravely ill — the commission recommends that available health-care facilities be expanded to include more hospices and outpatient care centres. More trained counsellors, nurses and social workers will be needed, and more emphasis placed on helping minorities.

Nearly half of the commission's recommendations were on basic research and drug development. Specific recommendations include the establishment of regional HIV research centres, a computer network to centralize AIDS research

information and new administrative regulations to force the administration to deal rapidly with requests from health agencies.

The commission proposes significant changes for the National Institutes of Health (NIH). Research grants would be restructured to cover longer periods, and the requirement to return money not spent by the end of the grant period would be dropped. Prizes, higher pay-scales and more training programmes would be created at NIH to draw in and retain the best scientific talents. To handle the increased burden, money would be authorized for constructing and equipping new laboratories on the NIH campus, something that has not been done since the late 1960s. NIH's AIDS reagent repository programme would also be a priority.

The commission recommends that drug development for AIDS should continue to focus on screening existing drugs for antiviral activity, but that more money should be spent on studying what makes a drug effective against HIV. The number of drug reviewers at the US Food and Drug Administration who handle HIV-related products would double to process the increased number of new drug applications. To eliminate the ethically troubling practice of administering placebos in clinical trials of AIDS drugs, an "historic control" database would be created. The commission also recommends clinical trials for studying the effects of proposed AIDS drugs in women, infants and intravenous drug users.

The total \$1,300 million for AIDS covered in President Reagan's budget request for 1989 (see *Nature* 331, 646; 1988) will not be enough to pay for these sweeping recommendations. But Congress is likely to embrace many of the suggestions, and its version of the budget may reflect this enthusiasm.

The commission will continue to hold hearings on AIDS-related topics until June. Forthcoming subjects include AIDS prevention and education, discrimination and ethical issues, testing and confidentiality, safety in the workplace, legal concerns, health-care financing and international cooperation.

Carol Ezzell

Superconductivity warning

ON pages 55–58 of this issue, the preparation is described of a novel superconductor which includes thallium in its composition. Note that thallium is a toxic substance and appropriate precautions should be taken in its use. □

Information body

PROPOSALS to create a new organization to provide Britain's parliamentarians with "unbiased, high-grade information" on issues of science and technology have moved a step closer to reality. The all-party Parliamentary and Scientific Committee, an unofficial body comprising members of both houses of Parliament, academics and industrialists agreed last week to establish a Parliamentary Science and Technology Foundation. The idea for such a body was first announced late last year (see *Nature* 330, 685; 1987).

The organization, which will employ three or four people, will be financed mainly by industry, the government having declined to authorize any money from Parliament's own funds. The body is being set up by senior Conservative backbenchers, unhappy with the present inadequate arrangements for informing rank-and-file members of parliament on topical scientific issues. Timothy Raison, also a Conservative, complained that members of the committee felt its brief was too broad to allow sufficient consideration of the state of British science. He suggested the creation of a new select committee to consider science separately. A more likely option, however, would be to allow the existing committee to set up a scientific subcommittee. S.H.

Britain looks to EC

BRITAIN'S five research councils have set up a joint office in Brussels to improve links with the European Commission and to learn how to tap the commission's research funds more efficiently. The five councils will share the costs of running the office. The research councils have had difficulty in learning about relevant European research programmes at the right time. S.H.

Superconductor links

AGREEMENT was reached last week on measures for improving European co-operation on research into high-temperature superconductivity. Representatives from the research-funding bodies of Britain (SERC), France (CNRS), Italy (CNR) and West Germany (DFG and MPG) met in Paris to discuss ways of implementing a more united approach to the research in order to compete effectively with Japan and the United States. The measures agreed include the free exchange of information on national research priorities and funding arrangements; the exchange of information on facilities and personnel; provision of funding for exchange of post-docs and provision of visiting fellowships; the establishment of a mechanism for disseminating new results before publication; and a method of providing a European basis for the refereeing of major research proposals. S.H.

West German AIDS research programme still in limbo

Munich

ENTERING the fifth year of a national research programme on AIDS (acquired immune deficiency syndrome), West Germany is eager to be seen doing its part in the worldwide effort to combat the disease.

But, despite a proclamation by Chancellor Helmut Kohl that "no reasonable project will fail for lack of money", West Germany has yet to make a mark with its programme. The few research groups with an international reputation are scattered geographically, and co-ordination of the programme is loose.

Now the Federal research and technology ministry (BMFT) is hoping to remedy this state of affairs with a number of "special collaborative programmes" modelled on the *Sonderforschungsbereichen* of the DFG (Deutsche Forschungsgemeinschaft). A large group of Munich researchers has already applied for funds under this plan, while a group of about 50 researchers in Göttingen, Hannover and elsewhere is working on an application for DM 2.5 million.

Part of the explanation for the slow start is that AIDS came later to West Germany than to the United States. In 1983, when BMFT first allocated special funds for AIDS research, there were only 10 cases. Even now, there are fewer than 2,000. Since 1983, BMFT has spent just DM40 million (about \$23 million) for basic research, compared with the annual US AIDS research budget of more than \$900

million.

The root of the problem lies with the researchers themselves. According to Hermann Wagner, a professor at the University of Ulm and head of the panel appointed by BMFT to judge the new proposals, "there are not enough good applications" for the money available.

Nobody claims that West German researchers lack talent or interest. The shortage of strong proposals seems to stem from factors such as these:

- **Structure.** BMFT, which funds 90 per cent of basic AIDS research, has departed from precedents elsewhere by not appointing a single person to supervise its AIDS programme and to identify new projects. The role played in the United States by Anthony Fauci of the National Institute of Allergy and Infectious Diseases of the National Institutes of Health is cited by many researchers as a model of what is needed here.

Instead, BMFT has appointed a committee to oversee its projects which, located in the Federal Health Office (*Bundesgesundheitsamt* or BGA) in West Berlin, is still feeling its way. In any event, BGA has much less than influence than the National Institutes of Health in the United States.

Within the West German university system, it also seems difficult to channel the funds available to the young, able researchers most likely to want to work on AIDS, partly because research grants from BMFT or the DFG do not include

"basic equipment" or laboratory space, all of which must be provided by the university, which then also has to pay much of the overhead costs of equipment.

One bizarre consequence of this arrangement in all research fields is that universities do not encourage outside research support for fear that their own costs will escalate. This problem ultimately traces back to the sovereignty of the *Länder* (states), not the federal government, over universities.

- **Culture.** The zeal to work on AIDS evident in the United States is missing in West Germany. One group of researchers returned from the United States impressed that finding a treatment or a vaccine for AIDS had become a kind of "national mission". But DFG President Hubert Markl rejects the "big mission" strategy for West Germany, saying that each country has its own best way of attacking problems. Although the "bandwagon approach" had been very successful in the United States, he says that in West Germany, "you can marshal more support by offering a base for future research". Markl admits that, so far, this approach has not produced "research at the cutting edge".

- **History.** West Germany is hampered by the lack of a centre for virology research. Ironically, there used to be one, a Max Planck Institute at Tübingen, but that was gradually reoriented in the 1970s to its current focus on developmental biology.

Many of those now working on AIDS cut their teeth at Tübingen under the virologist Werner Schäfer. Thus Gerhard Hunsmann, at the German Primate Centre at Göttingen, is supported by BMFT and European Community funds in a search for a primate model for AIDS other than the chimpanzee. Reinhard Kurth, at the Paul Ehrlich Institute in Frankfurt, is collaborating in the study of simian viruses with groups in the United States in a search for clues to a vaccine. Other Schäfer alumni are in West Berlin, Heidelberg and the United States, while there are also internationally competitive groups of virologists at Munich and Erlangen.

BMFT hopes that its new scheme will help overcome geographic separation and generally stimulate virology and related research, but some researchers are sceptical, noting that the necessary structural changes in the university system will not happen, as one put it, "unless we have a revolution or people are dying in the streets".

But others are more optimistic, saying that BMFT and the BGA are doing reasonably well, given that they started late. They point to the reputation of West German researchers in immunology, electron microscopy and related fields. But only time will tell how effective BMFT initiative in AIDS research will be.

Steven Dickman

Lowering the high risks of bad blood

Washington

A JUST-released study estimating the risk of contracting AIDS (acquired immune deficiency syndrome) from a blood transfusion at about one in 40,000 is already out of date, according to Gerald Sandler, deputy director of the American Red Cross. New screening procedures ensure that the risk is much lower, he says.

The study, published in last week's *New England Journal of Medicine* (318, 473; 1988), documents 13 cases where transfusion recipients became infected after receiving blood that had passed routine screening tests. The donors were later found to be positive for HIV antibodies, suggesting that they had donated blood after infection but before their antibody responses had developed to detectable levels. Five of the seven donors had been homosexually active and one had had sexual contact with an intravenous drug user.

Estimates of risk based on this study are too high, according to Sandler, because

new procedures help to exclude high-risk groups from donating blood and because antibody-screening tests are now more sensitive. The estimate was in any case pessimistic because it was a worst-case analysis, with data collected in high-risk areas.

Since the study was conducted, new procedures allow active homosexuals and others at risk to specify that their blood donation be used for experiment rather than transfusion.

But voluntary self-exclusion is not the only approach being pursued. Commercially available ELISA (enzyme-linked immunosorbent assay) kits for antibody detection are now more sensitive and have increased the chances of early detection. Tests for the presence of viral antigen are also in the pipeline and may turn out to be even more effective. The American Association of Blood Banks says its members will quickly make use of any new advances.

Penelope Austin

US health insurance companies screen applicants for AIDS

Washington

US health insurance companies are now routinely screening individual health insurance applicants for AIDS (acquired immune deficiency syndrome), according to a new report* from the congressional Office of Technology Assessment (OTA). The change has occurred without a national consensus on how those refused insurance should be helped, or on how an individual's level of risk of exposure to AIDS can be fairly assessed.

The number of people seeking individual health insurance is less than 10 per cent of the total insured population, and of this percentage half are elderly people who are unlikely to contract AIDS. It is the remaining half, made up of around 5 million individuals, that is causing concern to companies specializing in individual insurance. Most Americans are covered under group health insurance plans at work. These plans apply to sufficiently large numbers of people to cope with the few who will develop AIDS.

The insurance companies say that they must protect themselves financially by refusing insurance to those at risk from AIDS, just as they refuse insurance to those likely to suffer from cancer, coronary heart disease and mental illness.

But ethical problems remain over assessing a person's risk from AIDS, with homosexuals' activist groups arguing that homosexuality, which can place a person in a higher risk group for AIDS, cannot be regarded in the same way as smoking, which places someone in a higher risk group for cancer.

California and Washington, DC, have gone to the extreme of passing laws prohibiting the use of AIDS tests by insurers. Washington, DC, also bars companies from asking for medical information that might be used to assess a person's exposure to the AIDS virus. Insurance companies object strongly to these restrictions.

Most insurance companies do, however, agree that sexual orientation should not be a factor in underwriting. The National Association of Insurance Commissioners specifically advises state governments to forbid the practice. But according to the OTA report, 30 per cent of insurance companies take sexual orientation into account. Other companies ask applicants if they have suffered from sexually transmitted diseases common among promiscuous homosexuals. Tests for exposure to the HIV virus which causes AIDS are becoming commonplace. Half the insurance companies routinely test for HIV antibodies.

As the impact of the AIDS epidemic

grows, companies insuring individuals will become stricter, according to the report. HIV-antibody testing will be expanded and more AIDS-related questions added to applications. Companies are also more likely to examine AIDS cases to see if they can be linked to a pre-existing condition that would reduce liability. An insurance association survey has already shown that most AIDS-related claims are in the first two years after coverage begins, suggesting that individuals suspected that they were infected when they purchased insurance.

The Public Health Service has estimated that the cost of health care for AIDS patients will rise to between \$8,000 million and \$16,000 million dollars by 1991. As the epidemic spreads, there will be increasing pressure on state schemes which provide help for those who cannot obtain personal insurance. These rely on contributions taken from insurance company income. Insurance companies are calling for federal legislation to spread the burden wider.

Alun Anderson

*AIDS and Health Insurance: An OTA survey, February 1988

Easier access to UK defence labs

London

THE British government this week announced a new scheme to allow civil industry increased access to four Ministry of Defence (MoD) research establishments. The Civil Industrial Access Scheme (CIAS) is a joint initiative of the MoD and the Department of Trade and Industry. Industrial users of the establishments will be charged market rates for the use of facilities and for consultations with MoD staff.

Launching the scheme, trade minister John Butcher said he was confident that the venture would help strengthen the research arm of British industry. Butcher expects the scheme to be of particular interest to small- and medium-sized companies that cannot afford to invest in expensive equipment and facilities.

CIAS is part of the MoD's drive in recent years to exploit more fully research it undertakes. The establishments already carry out some civil work, mostly in aerospace and electronics — commercial spin-offs from defence work are exploited by a company recently set up by the MoD, Defence Technology Enterprises Ltd.

The new scheme will operate in the Admiralty Research Establishment, Royal Aircraft Establishment, Royal Armament Research and Development Establishment and the Royal Signals and Radar Establishment. Simon Hadlington

New superconductor magnets

Washington

THE new high critical temperature (T_c) ceramic superconductors could, in principle, be used to build the magnets for the Superconducting Super Collider (SSC), but a US Department of Energy (DoE) report* concludes that it will take at least 12 years to overcome the "formidable materials science and engineering problems" of the new materials.

The flurry of excitement last year over the new class of superconducting compounds prompted DoE to investigate their use as SSC magnets. In requesting the investigation, Alvin Trivelpiece, then director of the Office of Energy Research, admitted that the task was not easy, "considering the rapidity with which the science is developing", a point driven home by the discovery of bismuth-calcium superconductors (see *Nature* 331, 377; 1988).

A panel chaired by Albert Narath of AT&T Bell Laboratories has now determined that further development of the rare-earth oxides would achieve the 10^5 A cm⁻² critical current and 7 T critical magnetic field required to operate the SSC.

But the stringent performance requirements of accelerator magnets demand huge strides in materials processing before the new materials could even be considered for use in SSC. Under "optimistic conditions", the panel reckons it would take 4 years to develop processes to make continuous lengths of 5–10- μ m diameter superconducting wire with appropriate mechanical properties. Another 8 years would be needed to develop winding techniques for building prototype magnets.

It is also not simply a matter of exchanging yttrium-barium-copper oxide superconducting wire for the niobium-titanium wire used in current designs. The new materials have higher heat capacity, making it necessary to add active magnet protection in case of accidental loss of superconducting properties.

Other design features would have to be modified to use liquid-nitrogen cooling for SSC. Operating at 77 K, rather than 4 K, requires higher vacuums and larger magnet aperture modifications that could offset any savings from operating at the higher temperature.

Experience with niobium-tin superconductors underscores the trade-offs between materials with desirable electrical and magnetic properties, but poor mechanical properties. Niobium-tin wires have been available for 20 years, but are too brittle.

Joseph Palca

*Report of the Basic Energy Sciences Advisory Committee panel on high- T_c superconducting magnet applications in particle physics, US Department of Energy, Office of Energy Research, Washington, DC, 1987.

Scientists involved in Armenian environmental demonstrations

London

SCIENTISTS and cultural workers helped coordinate the Armenian demonstrations in Erevan last week, according to Sergei Grigoryants, editor of the unofficial Moscow journal *Glasnost*. Grigoryants, himself half-Armenian, managed to visit Erevan for two days during the demonstrations, and brought out a video film which confirmed both the size of the demonstrations (several hundred thousand people) and the good order that the organizing committee kept.

The identity of these scholars has not yet emerged, but they are probably among the 350 Armenian academics and intellectuals who in March 1986 launched a major environmental campaign in Armenia with their open letter to Mr Gorbachev protesting against growing levels of pollution in the Armenian republic. Throughout the Soviet Union, the policy of *glasnost* has generated many semi-official "citizens' initiatives" in defence of the environment, and in almost every case, many of those most vocal about existing or impending environmental hazards are equally active in "national" campaigns which the authorities find less acceptable. Even in the Russian SFSR, the campaign against the diversion of the Siberian rivers was spearheaded by the group of writers known collectively as the "villagers" who stress the importance of Russian traditions and values and who saw the diversion as a threat to irreplaceable Russian archaeological treasures. In the non-Russian republics, the link is still more strongly marked. Many Estonians, for example, see the environmental threat from the development of the phosphorite industry and the ethnic threat from the influx of Russian workers (which could make the Estonians a minority in their own republic) as two aspects of one problem.

In the case of Armenia and the "national" demonstrations calling for the transfer of the Nagorno-Karabakh autonomous region from Azerbaijan to Armenia, the original impetus seems to have come directly from an ecological protest. On 17 October 1987, there was a protest meeting outside the Nairit calcium carbide plant in Erevan, whose managers had persistently and flagrantly ignored the anti-pollution laws. This demonstration had the approval of the local party authorities, and took place peacefully. Its success, however, inspired about a thousand of the participants to stage a second demonstration the next day, demanding the restoration to Armenia of Nagorno-Karabakh. This time, the demonstrators, who had hoped to march to the headquarters of the Armenian Communist

party, were dispersed by the police first.

In spite of the attempted demonstration in October, last week's events, first in Nagorno-Karabakh itself and then in Erevan, seem to have taken Moscow by surprise. Official commentators gave widely varying estimates of the ethnic composition of the Nagorno-Karabakh region, ranging from "a predominantly Armenian population" (TASS) to "18 per cent Armenian" (Moscow radio world service). (Western estimates put it at around 75 per cent Armenian.) The quality of life in the region also seems unclear: the deputy chairman of Azerbaijan's Council of Ministers, Ayaz Mutabilov, said that in many fields of the economy and culture, Nagorno-Karabakh has "higher indicators than the average for Azerbaijan", whereas Mr Gorbachev, in his message to the peoples of Armenia and Azerbaijan, spoke of the many "shortcomings and difficulties" which have, he said, accumulated in the region.

The Armenian demonstrators have now announced a month's moratorium. The problem for the Soviet government is, however, wider than that of perhaps redrawing the boundaries of the Armenian and Azerbaijan republics. Ethnic consciousness is on the increase everywhere in the Soviet Union, not only in traditional centres of national dissent such as the Ukrainian and Lithuanian republics, but also, for example, in Uzbekistan (centred on the Uzbek writers' campaign to save the Aral Sea), and in Byelorussia, where, after decades of ethnic inertia, a republic-wide network of informal groups has sprung up, focused on the native language, the "blank spots" of history and environmental problems. Ecological concern, once seen as one of the "safest" topics for *glasnost*, is taking on "nationalist" overtones.

The Soviet authorities have so far simply issued warnings about what are and are not suitable subjects for protest, with the environmental still in the "permitted" category. In Czechoslovakia, however, the environment is seen as having a "political dimension", and a legally constituted environmental group in Bratislava, called Nahlas, has recently come under sharp criticism from the party for having compiled an extensive report on the state of the city's environment on the grounds that such information could play into the hands of anti-socialists abroad. The members of Nahlas (the name means "Speak up!") consider this criticism unfounded and, it is understood, are hoping to make their case known to like-minded groups of environmentalists within the Soviet Union.

Vera Rich

Outlook poor for agriculture

London

AGRICULTURAL research in Britain is destined to remain inadequately financed for at least the next five years, despite increasing investment from the private sector. In its 1988 corporate plan, published this week, the Agricultural and Food Research Council (AFRC) predicts that its total income will increase annually by around 3 per cent for the next five years, which is insufficient to meet rising costs.

AFRC, which has recently undergone a period of radical restructuring, has borne the brunt of government cutbacks in research funding over the past five years. Since 1983, AFRC has shed around 25 per cent of its workforce, one-third of those by compulsory redundancy, to its present total of 4,881. This number is expected to decline further to 4,540 by 1990-91, before increasing slightly to 4,680 by 1992-93, mainly due to an increasing proportion of period appointments supported by outside industry.

Funding from government has declined by 20 per cent in real terms since 1983-84 to the present level, including commissioned research from other government departments, of £98 million. Contracts from the Ministry of Agriculture, Fisheries and Food (MAFF) amount to around £45 million annually. But a cut in MAFF's research budget will result in a decrease of £4 million over the next two years for AFRC.

Income from commercial sources and other government departments presently stands at £14 million, with a projected increase to £22 million by 1992-93, when it will amount to 20 per cent of the council's total income.

The council will increase its support of research in universities, at present £7.5 million. After 1988-89, the level of grants committed will be increased by 4 per cent annually. Until now, however, AFRC has maintained its support of university research at the expense of the budgets of its own institutes. The present corporate plan warns that from now on university support could be limited.

Much of the university support is intended eventually to be channelled through interdisciplinary research centres. Bids are now being invited from institutions to host a centre in one of eight topics of strategic importance.

In the institutes, programmes will be reshaped over the next few years, with greater emphasis given to molecular biology and genetics in plant sciences and animal physiology, horticulture and engineering. Reductions will be made in support given to aspects of arable crops, grassland and animal production, and animal health.

Simon Hadlington

Inquest verdict on Met. Office failure to predict storms

London

THE inquest into the failure of the British Meteorological Office to predict the severity of the storm that killed 19 people on 15–16 October last year has not identified a culprit — unless it is the British Treasury. The reports of two separate inquiries into the affair were published by the Ministry of Defence last week.

The independent report by Sir Peter Swinnerton-Dyer (Chairman of the University Grants Committee) and Professor Robert Pearce (University of Reading), commissioned by the ministry after the Met. Office had mounted an internal inquiry, concluded:

- British weather forecasters are inadequately trained.
- Senior forecasters have anomalously low status within the Met. Office hierarchy.
- Cuts in government support of the Met. Office have resulted in minimum staffing levels and inadequate computer power.

The report of the internal inquiry set up by Dr John Houghton, director-general of the Met. Office, says that paucity of data coverage gave an inadequate description of the storm and the atmospheric structure in its vicinity. This report calls for an improved observational network, further

development of the forecasting models used and a review of procedures for handling warnings of severe weather.

The government says it accepts the recommendations of both reports, but that they will have to be carried out within the Met. Office's budget of some £105 million.

Swinnerton-Dyer and Pearce conclude that there were two reasons for the inadequate public warning last October. First, the computer forecasts available on 15 October disagreed, with the result that all forecasts underestimated the strength of the storm and, second, the duty forecasters

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Satellite picture of the storm at 08.19 GMT on 16 October 1987.

University of Dundee Satellite Station

UK environmental research remains against the ropes

London

BRITAIN'S contribution to environmental research of national and global importance is being eroded by starvation of public money. So much is clear from the corporate plan of the Natural Environment Research Council (NERC), published this week.

In real terms, NERC income from the science budget, excluding money for the British Antarctic Survey, is expected to decline by £2 million by 1990–91 to around £61 million. Income from commissioned research has failed to meet NERC's expectations and has forced the council to rethink its priorities. Last year's corporate plan had expected income from commissioned research to be at around £34 million by 1988–89. The revised figure is nearer to £29 million. NERC has already announced that it may have to introduce compulsory redundancies for the first time as part of a series of cost-cutting measures.

NERC says that to meet its defined strategic objectives it would require £150 million by 1989–90. Current prospects fall short of this by £22 million, with the gap increasing to £35 million by 1993–94.

NERC's revised plans will result in a significant reduction in its support for university research. The council had originally intended to increase expenditure in universities by £1.9 million in 1988–89 from its existing level of £13.4 million. Now it has been forced to cut £0.7 million.

From a total staff of 3,263 at the beginning of 1983, NERC employed 2,733 by the end of 1987, a reduction of 16 per cent. The 1987–88 forecast of shedding 100 posts was not achieved, largely because of a drop in the numbers of people volunteering to leave prematurely. The council's present plan identifies the following programmes that will be foregone or restricted if it does not receive the necessary annual income.

- £10 million for geological surveying.
- £5 million for interdisciplinary research centres. Specific proposals have been made for centres, including population biology, deep crystal studies and environmental microbiology.
- £7 million for strategic programmes in agriculture, environment, the North Sea project and the biogeochemical ocean flux study.

Simon Hadlington

failed to recognize a situation in which computer models were likely to underestimate the strength of the winds. But this report exonerates the duty forecasters, saying that there was "considerable unlikelihood in the circumstances".

Last October's storm seems to have been a consequence of unusual meteorological conditions. The report links hurricane Floyd, which had developed in the Caribbean by 11 October, and the hurricane-force winds reaching south-east England on the night of 15 October.

The British storm was an intense mid-latitude cyclone, the energy of which derives from the contrast between warm air from subtropical regions and cold air from polar regions. The Earth's rotation means that the warm air moves upwards and ahead of the storm towards the pole and the cold air sinks towards the Equator behind it. Belts of strong upper and low-level winds (low-level jet) were associated with these regions of ascent and descent.

The gusts of hurricane force at the surface were the consequence of turbulent eddies in the lower atmosphere bringing air down from near the low-level jet. An upper-level jet streak, a region of 150 m.p.h. winds at a height of about 10 km, probably originating from hurricane Floyd, arrived off the Bay of Biscay and was a primary factor in causing the rapid development and deepening of the depression.

The report acknowledges that the observational network is poor to the south-west of Britain, where the storm originated, when computer models can give inadequate forecasts. It says that on this occasion the forecasters "largely accepted" the predictions of the model and that they "were either not aware of or underestimated the importance of the low-level jet".

The Met. Office has two forecasting models of low and high resolution ('coarse-mesh' and 'fine-mesh'). As reported earlier in *Nature* (329, 750; 1987), the forecasters were faced with conflicting forecasts. Although the coarse-mesh model is regarded as inferior, the fact that it is run later than the fine-mesh model meant that on 15 October it was crucially based on more up-to-date observations.

On the performance of the French meteorological services in forecasting the storm, the independent report says that the French forecasters displayed a "better appreciation of the nature of the phenomenon they were dealing with", perhaps because of the deeper training they receive. Also, the computer used by the Met. Office is less powerful than the French, with the result that the French service can run a fine-mesh model of twice the resolution. This is an "unnecessary handicap" says the report, and the authors are "relieved" that the Met. Office is to have a new ETA 10 supercomputer soon, even if the cost will have to be met through internal economies. **Philip Lloyd**

US Army to dispose of ageing chemical weapons on-site

Washington

AFTER years of debate, the US Army finally conceded last week that disposing of its ageing stockpile of single-agent, or unitary, chemical weapons at the sites where they are stored is preferable to transporting them to a single national incineration facility. Critics had raised the spectre of an accident en route with catastrophic consequences.

The Army has wanted since the early 1980s to exchange its unitary chemical weapons for newer binary weapons because the unitary stockpiles are ageing and in danger of leaking. Binary weapons are intrinsically safer, because they are made from two chemical agents that are individually safe and only become toxic when mixed.

Although the precise size of the chemical weapons stockpiles is classified, 30,000 tons is a reasonable estimate. Approximately half that amount is mustard gas, a skin-blister agent, and the other half is nerve gas. Approximately 65 per cent of the chemical inventory is stored in one-ton steel containers, the rest in weapons: projectiles, cartridges, mines and rockets containing propellants.

In 1985, Congress authorized the Army to begin acquiring the materials it needed for the new weapons, but delayed plans to assemble them until 1987. At the same time, Congress required the Army to develop plans for destroying its unitary stockpiles. By law, the "demilitarization" process should be complete by 1994.

But destroying the weapons has been a thorny problem, for both political and technical reasons. The Army's initial plan for a single incinerator met with opposition from a Congress worried about an accident during transport of the weapons from the eight storage arsenals.

On-site incineration will also be politically difficult. States and local communities may have grown accustomed to being the resting place for weapons, but being their burial ground is another issue.

The Army also ran into engineering problems in demilitarizing its stockpiles. A pilot incinerator at the Tooele Army Depot in Utah had to be shut down last year after it developed a leak. An alternative incineration method called "cryofracture" — where weapons are frozen and then handled by robots — has not yet been sufficiently successful to be relied upon, although the Army will continue to develop the process. There is also the problem that incineration itself will generate toxic by-products. Under the current plan, the weapons will be incinerated after partial disassembly. The inert ash will then be transported to an as-yet undetermined burial site.

But disposing of the weapons is expensive and time-consuming. It will cost between \$2,000 million and \$3,000 million, and even in the unlikely event that there are no further political or technical snags, the Army estimates that it cannot

finish the project before 1996.

There are also no plans as yet for dealing with the weapons currently stored in West Germany. The German government has indicated that it would like the weapons removed.

If the Army has been forced to delay its timetable for destroying old weapons, at least it was able to begin building new ones on schedule. The first 155-mm artillery shells using binary nerve gas came off the assembly line last December.

Impetus for developing techniques for destroying chemical weapons also comes from improved prospects for a treaty that would ban all chemical weapons. At present the Geneva Protocol of 1925 bans the use of chemical weapons in war, but does not prohibit countries from producing them. The United States has vowed never to be the first to use such weapons in combat. A draft chemical weapons treaty is being developed in multilateral arms control talks in Geneva, and bilateral technical discussions are also under way with the Soviet Union. Any treaty would undoubtedly call for the destruction of existing stockpiles.

Joseph Palca

Call for international treaty to protect genetic diversity

San Jose, Costa Rica

DEMANDS for an international treaty to protect the world's genetic diversity were made at the meeting of the International Union for the Conservation of Nature and Natural Resources (IUCN), an organization of scientists, governments and non-governmental bodies, held last month in Costa Rica.

The draft treaty won the backing of the 1,000 or so participants at the conference. It proposes protecting sites with high biological diversity using funds generated by a levy on commercial and industrial users of genetic material. It would also give countries the right to impose "reasonable charges" on the use of genetic material, a condition designed to please developing countries which want compensation for banning development in species-rich areas.

The complicated issue of compensation has been under discussion by the United Nations Food and Agriculture Organization for several years, without clear result. The IUCN suggests charging profit-making individual companies or industry-wide groups of companies either as a fixed sum or as a percentage of profits.

Guaranteed rights of access to genetic material for pure or applied research and for genetic engineering are also provided for in the draft treaty, although questions of whether countries actually own endemic species are not addressed.

Scientists in the IUCN argue that

genetic resources are better protected *in situ*, by preserving species in their natural habitat to protect the full range of genetic variability, an advantage not shared by conservation in gene banks, botanic gardens or zoos. Thus the emphasis is on conservation of habitats, and on obtaining funds for that purpose.

Participants also discussed preserving genetic diversity in developing countries through a "debt for conservation" plan, where banks are asked to donate bad debt to developing countries for use in conservation projects. Such a scheme is already under way in Bolivia (see *Nature* 328, 373; 1987).

At least two international banks are seriously considering donating more than \$1 million to Costa Rica for conservation, according to Costa Rica's Minister of Natural Resources, Dr Alvaro Umana. The Costa Rican Central Bank has agreed to pay 75 cents on the dollar for debt notes, in local currency, if the money is put into conservation projects. Costa Rica's government hopes to use the donations to restore more than 70,000 hectares of degraded pastureland to tropical dry forest in the Guancaste region.

Costa Rica's president, Dr Oscar Arias, winner of the Nobel Peace prize, told the conference that preserving genetic diversity is especially important for biotechnology which can solve some of the most urgent problems of developing countries.

Kathy Johnston

Guide to Europe

London

THE Science and Engineering Policy Studies Unit of the Royal Society has produced a guide entitled *European Collaboration in Science and Technology* which gives details of one hundred schemes for promoting research collaboration within Europe. The guide costs £25 and is available from the publication sales department of the Royal Society, 6 Carlton House Terrace, London SW1Y 5AG, UK.

Simon Hadlington

Better outlook next decade for young West German researchers

Bonn

Ask a young West German academic who will pay his salary next year, and he is likely to look to the heavens. The job situation in many fields has been bleak. But according to the Wissenschaftsrat, West Germany's influential science council, that will all change in the 1990s.

Wissenschaftsrat's chairman, Kurt Kochsiek, held a press conference here on 2 February to "give courage" to young researchers considering careers in academic life. Kochsiek cited a new council survey which showed that the turnover of tenured professors in most university faculties will return to normal levels of 4-5 per cent a year by the mid-1990s. Turnover has recently been less than one per cent a year in some areas.

But Kochsiek was the first to acknowledge that the calculations depend on the number of tenured positions, now about 16,000, remaining stable. Hubert Markl, president of the Deutsche Forschungsgemeinschaft (DFG), was quick to agree. The DFG funds much of the basic research at universities in West Germany.

Kochsiek said that there is already a healthy turnover in some fields, such as engineering, computer science and Protestant theology. By 1995, there will be good opportunities in veterinary sciences, agricultural sciences and architecture. Humanities and social sciences will remain problem areas.

The Wissenschaftsrat did not call for the creation of any new funding programmes to keep qualified young people in the universities until positions become available. There are already a number of such

programmes. The Heisenberg programme, for example, gives multi-year support to as many as 60 young researchers each year who have already completed a *Habilitation*.

But Markl added that disgruntled academics are poisoning the atmosphere. "It's tremendously discouraging when the students see someone who has worked hard but still doesn't get a job".



Kurt Kochsiek — looking to a brighter future.

An even bigger problem will be the threat of cuts in universities as the number of students starts to drop in the 1990s. Some of the *Länder*, which run the universities, have already begun to cut back despite overflowing laboratories and lecture halls. In Markl's view, even small cuts would be "disproportionately damaging" because of the negative message they would send to good students. **Steven Dickman**

Indian ballistic missiles on the way

New Delhi

INDIA successfully test-fired its home-produced surface-to-surface missile Prithvi that can deliver a one-tonne warhead on a target 250 kilometres away in less than 150 seconds. It is the forerunner of the intermediate-range ballistic missile (IRBM) that India is likely to test later this year.

Announcing Prithvi's launch from Sriharikota base, Prime Minister Rajiv Gandhi said that it belongs to a class of missiles produced only by the United States, Soviet Union, France and China. He said it was designed and built by scientists at the Defence Research Development Organisation (DRDO) and "no foreign knowhow or collaboration was involved".

Prithvi is powered by two gymbal-mounted engines burning high-specific-impulse liquid propellant. It is designed to leave the ground in 10 seconds after command, and

its entire flight is controlled by a strap-down gyro system coupled to an on-board computer. According to a DRDO spokesman, it can carry "different types of payload" and compared to other missiles of this class, "has the best warhead-to-weight ratio." Describing the test flight as "far beyond expectation," he said telemetry data showed a bullseye hit, with circular error of probability (CER) almost zero.

Prithvi is the second missile to come out of DRDO's five-year old integrated guided missile development programme. Trishul, a short-range (10 km) surface-to-air missile developed last year is soon to enter production. The ultimate aim is to develop an intercontinental ballistic missile.

The missile programme is no doubt worrying Pakistan where the latest launch is interpreted as Indian preparation for a nuclear delivery system. **K.S. Jayaraman**

New money for Caltech

Pasadena

THE California Institute of Technology (Caltech) is poised to give a team headed by Leroy Hood a spacious new facility for developing advanced DNA sequencing technology and exploring the uses of computer chips and neural networks for the analysis of sequence data. Such largesse is made possible by the new Beckman Institute to be located on the Caltech campus, where space and money will be available for exploratory research projects and interdisciplinary technology development.

The institute, financed by a \$50 million grant from the Arnold and Mabel Beckman Foundation and \$10 million in matching funds raised by Caltech, will generate no new faculty positions, but will instead consist of flexible space where existing faculty can set up 'resource centres' for technology development at the interfaces of chemistry, engineering, biology and medicine.

Jokingly dubbed "The Beckman Institute for Harebrained Research" by its director, Caltech chemistry professor Harry Gray, the institute will also provide small (\$50,000-\$100,000) start-up grants for interesting high-risk projects that may not be immediately fundable by government agencies but which promise important advances should they succeed.

Gray expects there to be four or five resource centres in the institute at any one time, each with 10-50 research staff and 5,000-10,000 square feet of space. Caltech plans to spend \$25 million on the 140,000-square-foot building, scheduled for completion in October 1989. That leaves \$35 million as an endowment for research support, which will provide about \$1.5 million in research funding each year. The institute's research costs are likely to be 10 times that amount, so most of the research will need to quickly become self-supporting with government grants.

The first round of proposals for resource centres was recently received by the six-member in-house faculty committee, and will be reviewed over the next few months. The 14 proposals include plans for ultrasensitive mass spectroscopy to be used for protein sequencing, and development of laser spectroscopy to study reactions on a femtosecond timescale, as well as Hood's facility.

Hood's proposal, frequently emphasized by Caltech during the fund-raising phase of the project, is almost certain to be among those chosen. It will require 10,000 square feet of space, says Hood, and at least \$2 million a year in operating funds, for which he is seeking support from the National Science Foundation.

Marcia Barinaga

Access to scientific journals

SIR—The advance of science depends upon continuous communication among its practitioners. Interchanges of new concepts and laboratory and field data occur at meetings, through personal contacts and through various forms of publication. But many members of the scientific community have only limited contacts with their colleagues, particularly in the developing world where hard currency, and hence travel and scientific journals and books, is difficult to come by. The lack of journals that publish a wide spectrum of important scientific articles, such as *Science* and *Nature*, is especially critical. Such publications are necessary stepping stones for innovative and substantial studies.

The lack of exposure to these journals by a large number of scientists became evident to me during a recent sabbatical at the Ruder Boskovic Institute in Yugoslavia. This distinguished institution, with more than eighty scientists and a large number of laboratories outside its central facility in Zagreb, has single subscriptions to both *Science* and *Nature*, which are kept under lock and key in the Zagreb Library. Personal subscriptions are out of the question. As a consequence, most investigators are not exposed to the journals in continuous and systematic ways. This situation is repeated over and over again in the developing world.

A second observation in Yugoslavia suggests an interesting solution to the problem. I visited a number of institutions and saw in their libraries a superabundance of periodicals from international agencies such as UNESCO, UNEP, FAO and WHO. Although some were substantial and relevant, the local scientists emphasized that others were of poor scientific quality and of little use. The scientific presses of these organizations are most productive of pages but often not of substance. In addition, many of these publications end up in bureaucratic offices, not in the hands of the scientists for whom they were intended.

A small percentage of the publication costs of these international agencies, if directed to the reproduction and distribution in adequate numbers of articles and reports from *Science* and *Nature* to the laboratories and scientific facilities of countries in the developing world, would constitute a major benefit. Clearly, such problems as copyright restrictions could be overcome through effective political action.

Who might bring about such action? The new director general of UNESCO will clearly have ample opportunities for innovative activity. Or perhaps a collusion of scientists of the developing world might push for such a venture. I suspect, how-

ever, that activity must be initiated from the top down, not from the bottom up.

EDWARD D. GOLDBERG

*Scripps Institute of Oceanography,
University of California, San Diego,
La Jolla,
California 92093, USA*

Science in Spain

SIR—Your supplement on "Science in Iberia" (*Nature* 324, 313–332; 1986) described most of the new aspects of Spanish scientific policy. The national plan for scientific research and technological development will allow up to 1 per cent of gross domestic product to be spent on research and development. This figure, although not large enough to be compared with that in other developed countries, represents some effort by the Spanish government and makes it possible to be reasonably optimistic. Nevertheless, the human factor was not taken into account in your report.

Most researchers believed that the reorganization of science in Spain would result both in an increase in the number of scientists and in a change of the traditional climate, which lacked social stimuli. New appointments have allowed young people to be recruited to the staff of the Consejo Superior de Investigaciones Científicas (CSIC), but it is generally agreed that senior people have hardly been promoted. On the other hand, 18 months after the law for the promotion and general coordination of scientific research and technological development was passed by parliament, salaries of researchers at both the CSIC and the universities continue to be shockingly low — about 30 per cent less than the average salaries of other comparable public employees. This fact was publicly conceded by CSIC's head, Professor Trillas, at a meeting in Madrid last December. The complete reorganization of the CSIC's personnel that is now being carried out would allow a substantial increase in salaries, which should represent only a small economic effort. But the Spanish authorities seem to have chosen to lose this historic opportunity and to ignore the advice of scientific associations and trade unions. This lack of receptiveness is discouraging to many Spanish scientists and could seriously disturb the development of current and future research projects. It would be a great pity if a government that has invested such a great effort in the renaissance of Spanish science were responsible for its failure.

ERNESTO GARCIA

*Centro de Investigaciones Biológicas,
CSIC,
Velázquez, 144,
28006 Madrid, Spain*

Germ-line quandary

SIR—I find it remarkable that your recent comment on germ-line genetic manipulation (*Nature* 331, 100; 1988) itself overlooks a very simple practical point, while complaining that "several important and ultimately practical questions have been ducked".

In considering manipulation of the "genomes of early embryos so as to avoid the possibility that unwanted genetic disease will emerge" the writer took as an example the X-linked disorder haemophilia. Clearly, a prerequisite for a 'correction' of the mutation by genetic manipulation will be the ability to diagnose the mutation in the early embryo in the first place. From our present knowledge, it seems that both the diagnosis and manipulation would need to be in a preimplantation embryo generated by *in vitro* fertilization. A carrier for haemophilia has a 50 per cent chance that her male embryo would not have the mutation and the current widely accepted clinical practice of *in vitro* fertilization generates several preimplantation embryos simultaneously. Thus our future mother-to-be and clinical molecular embryologist will have a stark choice; to be the first to transfer a genetically manipulated human embryo in the hope that the experiment will lead to the birth of a healthy baby, or to select another embryo that does not carry the haemophilia mutation. There seems little doubt from both a practical and an ethical standpoint that embryo selection will win hands down, when helping couples to avoid the transmission of serious genetic disease to their children.

Current legislation in the making recognizes the benefits of preimplantation diagnosis and embryo selection and, it is hoped, will allow, within strict limits, some research into early human embryology. To argue that genetic manipulation of early embryos should not be outlawed without giving any example of potential benefit that could not be better provided by embryo selection is particularly foolish at this time of public debate (and when the British government's white paper on the new embryology is to be debated in parliament). There are many who would wish to stop all embryo research, including studies designed to test the safety of embryo biopsy on which preimplantation diagnosis will depend. Vague claims of benefit without giving clear practical examples will only hasten "the drift into muddle" that your leading article sought to warn us of, and damage the credibility of clinical scientists in the process.

MARCUS PEMBREY

*Mothercare Department of Paediatric
Genetics,
Institute of Child Health,
30 Guildford Street,
London WC1N 1EH, UK*

Earthquakes and the Earth's rotation

A new calculation of the effects of past earthquakes on the speed of the Earth's rotation does not account for the observations, but raises important questions in geophysics.

THAT the speed of the Earth's rotation is not constant is, of course, familiar — although the world's observatories can still count on a little publicity for their annual announcements that the year just passed happened to be a few milliseconds longer or shorter than expected. Moreover, there is no shortage of explanations for seasonal and secular changes of the speed of rotation. Differential expansion of the atmosphere north and south of the Equator is one obvious cause of seasonal changes. Tidal dissipation and post-glacial changes in the distribution of material (melting ice followed by post-glacial rebound) have secular consequences, while earthquakes are usually also listed among the reasons why the moment of inertia of the Earth may be changed, but only qualitatively.

Now, mercifully, the issue has been made quantitative, by means of an interesting study by B. Fong Chao, of the Goddard Space Flight Center in Maryland, and Richard S. Gross, of the Los Alamos National Laboratory (*Geophys. J. R. astr. Soc.* **91**, 569; 1987). Unsurprisingly, the authors of this study find that earthquakes do indeed serve to make the Earth more compact, thus decreasing its moment of inertia and, because they leave total angular momentum unchanged, increasing the rotation speed and thus decreasing the length of the day (LOD), which is what would be expected.

But there is a substantial puzzle. The size of the calculated effect of all the substantial earthquakes (of magnitude greater than 5.5) between 1977 and 1985 is two orders of magnitude smaller than that expected from observation. Moreover, there seem to be systematic trends in the data suggesting that earthquakes do not occur randomly, but that there is a bias in favour of those events which, among other things, decrease the LOD. Quite properly, Chao and Gross say that this raises awkward questions about the interaction between the mechanism of earthquakes and other geophysical processes, those of plate tectonics for example.

Tackling such a complicated problem is naturally a complicated business, involving first a scheme for representing the gravitational potential of the Earth in terms of spherical harmonics and then of working out a way of calculating the changes of density distribution caused by individual earthquakes. Chao and Gross are able to relate the changes of the iner-

tial constants of the Earth to the seismic moment of an earthquake, and discover, with the approximations they use in passing, that the effects of separate earthquakes are cumulative — whence the feasibility of their calculation of the effects of 2,146 events over nearly a decade.

Apart from simplifications which are mathematical approximations, there is however one that could deserve much closer attention than it has been given. Chao and Gross work with a symmetrical model of the Earth, noting that departures from sphericity are only one part in 300 or thereabouts, and that the effect of these small variations on the calculation of other geophysical quantities appears to be small. But is there a chance that, in programmes of calculation from which interesting departures from randomness emerge, the non-spherical character of the Earth may be qualitatively important, perhaps helping to explain the large difference between the results of calculation and observation? Chao and Gross would be well within their rights to say that questions such as that are more easily asked than answered.

The outcome of the calculations as they are is disconcerting, but not surprisingly so. To the extent that earthquakes should change all the components of the Earth's inertial tensor, the off-axis components as well as the principal moments of inertia along the diagonal, earthquakes should be one of the causes of the measured drift of the position of the Earth's rotation axis, or the geographical position of the poles, which is now well measured by the laser ranging of the motion of Earth satellites.

In the event, the calculated motion of the poles is in almost the opposite direction to that observed, which merely shows that other causes of excitation are more important than earthquakes. But it also seems that the calculated cumulative effect of earthquakes has been to nudge the position of the pole towards 150°E ever since good seismic records were first kept at the beginning of the century. That is one of the non-random effects to which Chao and Gross rightly draw attention.

The calculated effect of earthquakes on the inertial constants of the Earth is equally striking, affecting the moment of inertia about the rotational axis but not the two components at right angles. This, of course, is the quantity most directly linked with LOD. The result, as expected, is a cumulative decrease of the moment of

inertia, but not a steady decrease: some earthquakes increase the moment of inertia. But the value of the rate of decrease works out at merely 0.1 microseconds a year, roughly one per cent of the measured rate of decrease. One conclusion is that the effects of earthquakes are swamped by those of other processes, another is that the calculations may underestimate what earthquakes do.

The importance of this calculation is that it provides a technique that may be applied, cumulatively, to events that have not yet occurred. It is not unreasonable to hope that, as the years go by, it may be possible to recognize in the short-term variations of the Earth's rotation speed the effects of major earthquakes, however great may be the other influences upon it. That, in the end, is how Chao and Gross will be tested.

Meanwhile, the questions remain of why the effect of earthquakes on the Earth's rotation should have the effect of predominantly decreasing the polar moment of inertia (and also of nudging the rotation pole in a specific direction). Chao and Gross are right to say that these questions point deeply into geophysics.

Even the question of why the Earth's matter is redistributed, as a consequence of an earthquake, so as to decrease the moment of inertia, is not as simple as it sounds. It is true that this is the direction in which the potential energy of the system as a whole will be decreased, but that does not give a good account of how the rocks about to rupture sense what the principle of least energy requires of them. The fact that some earthquakes increase the moment of inertia shows that the problem is not simply a static problem.

If there are preferred directions built into the problem, such as the tendency for earthquakes to nudge the rotation pole in a specific direction, the need for an explanation linking the causation of earthquakes with tectonic processes will be even more apparent. For while it requires no imagination to appreciate how earthquakes along, say, the San Andreas Fault, are driven by the well-logged relative motion of the Pacific plate, consistency in the direction in which the rotation pole is moved by earthquakes would suggest a much more global bias in favour of earthquakes with particular seismic moments. Chao and Gross ask if the phenomena point to some "behind the scenes" process not yet recognized.

John Maddox

Nonlinear dynamics

Observations of real solitons

Peter L. Christiansen

SOLITONS are localized nonlinear waves that can propagate and interact like particles. Theoretical studies show that phenomena such as water waves, light pulses in optical fibres, magnetic-flux quanta in superconducting devices and coherent excitations of biomolecules can be solitons. Computer simulations show that solitons can form in the presence of such realistic features as frictional loss mechanisms, external driving forces and thermal fluctuations. The solitons will exist under these circumstances for sufficiently long to be important features in the time evolution of the wave excitations. Experimental demonstrations of soliton dynamics, however, are still scarce. Therefore, two recent papers by Fujimaki, Nakajima and Sawada¹ and by Wu, Wheatley, Putterman and Rudnick² showing solitons in real systems are most noteworthy.

The work by Fujimaki *et al.* deals with collision of solitons on an electronic Josephson transmission line (JTL), 1.8 mm long, composed of a sequence of 31 discrete Josephson junctions (interleaved superconducting and insulating layers). In the continuum version of the JTL, the Josephson effect (superconducting electrons tunnelling through the insulating layers) results from the weak coupling between pairs of superconducting thin films. This overlap geometry is modelled very accurately by the sine-Gordon equation originally developed by particle physicists. In 1962, Perring and Skyrme³ showed that this nonlinear partial differential equation possesses solutions that they termed 'kink' and 'antikink', after

their shapes, which can propagate and interact with each other completely non-destructively, suffering only phase shifts as a result of the interaction.

In certain respects, Perring and Skyrme's paper foreshadows the pioneering work by Zabusky and Kruskal on the Korteweg-de Vries equation (originally formulated to describe shallow water

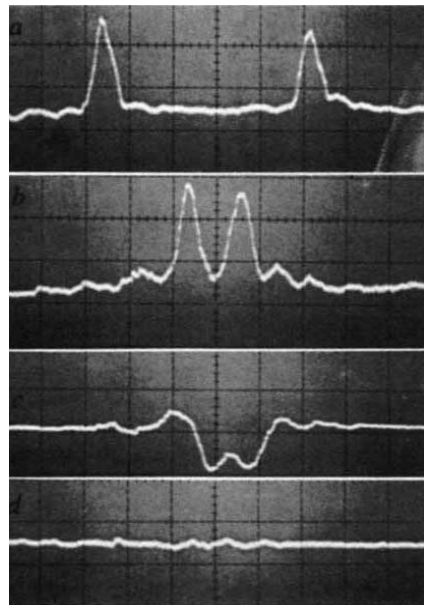


Fig. 2 Fluxon-antifluxon annihilation process observed *a* 0, *b* 8, *c* 16 and *d* 24 ps after the fluxons are launched. The width of the picture is 1.8 mm. (Courtesy of A. Fujimaki.)

waves), which introduced the soliton concept. The derivative of Perring and Skyrme's kinks corresponds to the solitons. An isolated kink solution to the sine-Gordon equation cannot be destroyed and is therefore a 'topological' soliton. In the JTL theory, these solitons carry magnetic-flux quanta (current vortices) and are termed fluxons. Correspondingly, negative magnetic flux quanta are carried by antifluxons. One consequence of the theory is that the fluxons behave as relativistic particles (their effective mass increases with increased velocity). In both papers^{3,4} the analysis was directly inspired by computational results.

The superconducting circuit, made using standard lead-based technology, consists of two fluxon generators, the JTL and a sampler to measure the flux in the JTL (Fig. 1); all are based on the Josephson effect. A fluxon is injected into the JTL when a control current I_c to the fluxon generator exceeds a critical value. A bias current I_b perpendicular to the JTL controls the velocity v of the fluxons (which is

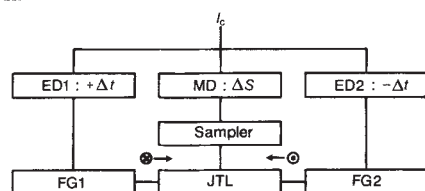


Fig. 1 Simplified block diagram of the fluxon experiment of Fujimaki *et al.* The fluxon and antifluxon (cross and dot, respectively) are injected into the Josephson transmission line (JTL, made of 31 Josephson junctions, each a $4 \times 4\text{-}\mu\text{m}^2$ sandwich of two superconducting films and one insulating) by the fluxon generators FG1 and FG2. Electronic delays ED1 and ED2 delay the fluxon and advance the antifluxon by Δt , effectively shifting the system $v\Delta t$ leftwards along the JTL. The sampler records the flux density at the midpoint of the JTL after a variable mechanical time delay (MD) of ΔS . The whole process is initiated when the control current I_c exceeds a critical value. (From ref. 1.)

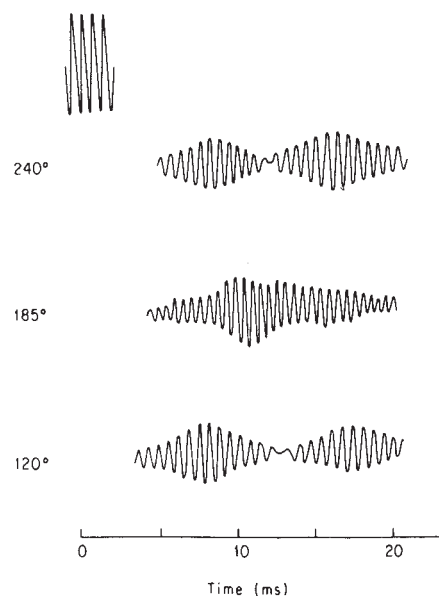


Fig. 3 Collision of clockwise and anticlockwise envelope solitons travelling around a thin cylindrical shell. The shape of the envelope persists because of the balance between amplitude dispersion and the nonlinearity of the elastic medium. Without this balance the wave packet would disperse. The top trace shows the excitation pulse. (From ref. 2.)

typically $4 \times 10^7 \text{ m s}^{-1}$), and must also exceed a minimum value. The sampler detects only at one central point, so that the flux across the whole JTL cannot be measured in one go.

Instead, a variable delay $2\Delta t$ can be introduced between the two fluxon generators, effectively moving the fluxon-antifluxon pair left or right by a distance $v\Delta t$. A single scan, such as shown in Fig. 2, is obtained by sweeping Δt . The time sequence shown in Fig. 2 is obtained by introducing a second delay ΔS before firing the sampler.

Fujimaki *et al.* thus observe the collision of a fluxon and an antifluxon, in which the two merge, interact and dissipate their energy via a 'breather mode' (the soliton and antifluxon bound together in a localized, lower-energy state) into linear oscillations ('radiation') which are finally absorbed by shunt resistances (Fig. 2). This agrees well with computational results and can also be understood in terms of fluxon perturbation theory⁵. In other cases, the fluxons can survive the collision. The fringes in the traces are non-solitonic effects inevitable in real versions of such idealized mathematical concepts. The authors also demonstrate the relativistic nature of the fluxons by showing that the product of the pulse height and the pulse width is independent of the fluxon velocity.

Wu *et al.*² observed envelope solitons in elastic solids for the first time. In contrast to topological solitons, envelope solitons are wave packets with a soliton-shaped envelope (Fig. 3). Earlier experiments with fluid surface waves have been

interpreted in terms of envelope solitons⁶. Wu *et al.*² excite flexural wave packets by an acoustic horn driver on a circular-cylindrical thin elastic shell. The waves propagate clockwise and anticlockwise around the cylinder. Their nonlinear interaction is seen by means of transducers mounted on the shell. The reason for using thin shells is that this system is very dispersive and has a high nonlinear response, making the generation of solitons feasible, and a large quality factor, so that they persist. Such a combination of parameters generally makes soliton formation in off-equilibrium (highly distorted) systems accessible to observation.

As in other quickly varying hyperbolic classical fields, the slow modulations of the flexural waves on the shell are described by the nonlinear Schrödinger equation. The parameters in the interacting envelope soliton solutions of this equation can be fitted perfectly to the observed data, thus implying the existence of envelope solitons.

Solitons are not of interest to physicists alone. Davydov⁷ has proposed application of envelope solitons to understand energy storage and transport in protein chains. Our computer studies⁸ of the quantum model of α -helical proteins show that the solitons have long lifetimes at biological temperatures. Unfortunately, experimental evidence of this mechanism remains to be obtained. □

1. Fujimaki, A., Nakajima, K. & Sawada, Y. *Phys. Rev. Lett.* **59**, 2895–2898 (1987).
2. Wu, J., Wheatley, J., Putterman, S. & Rudnick, I. *Phys. Rev. Lett.* **59**, 2744–2747 (1987).
3. Perring, J.K. & Skyrme, T.H.R. *Nucl. Phys.* **31**, 550 (1962).
4. Zabusky, N.J. & Kruskal, M.D. *Phys. Rev. Lett.* **15**, 240–243 (1965).
5. Pedersen, N.F., Samuelsen, M.R. & Welner, D. *Phys. Rev. B* **30**, 4057–4059 (1984).
6. Yuen, H.C. & Lake, B.M. *Phys. Fluids* **18**, 956–960 (1975).
7. Davydov, A.S. *J. theor. Biol.* **38**, 559–569 (1973).
8. Cruzeiro, L., Halding, J., Christiansen, P.L., Skovgaard, O. & Scott, A.C. *Phys. Rev. A* (in the press).

Peter L. Christiansen is at the Centre for Modeling Nonlinear Dynamics and Irreversible Thermodynamics (MIDIT), The Technical University of Denmark, DK-2800 Lyngby, Denmark.

Vaccine development

Antigenic hybrids of poliovirus

James M. Hogle

AT A recent meeting*, three research groups reported the successful construction of viable intertypic chimaeras (hybrids) of poliovirus. In each case, the infectious cloned complementary DNA encoding one of the three serotypes of the virus was modified so that the amino-acid sequence of one of the antigenic sites of one serotype is replaced with the analogous antigenic site from a different serotype. All three groups focused on a particular region of the virus capsid protein, VP1, that is known to be a dominant antigenic site of type 2 and type 3 poliovirus (see figure)¹. These studies have important implications for the design of novel vaccines for polio and related viruses, and on page 81 of this issue², Burke *et al.* present the first published report of such an antigenic chimaera.

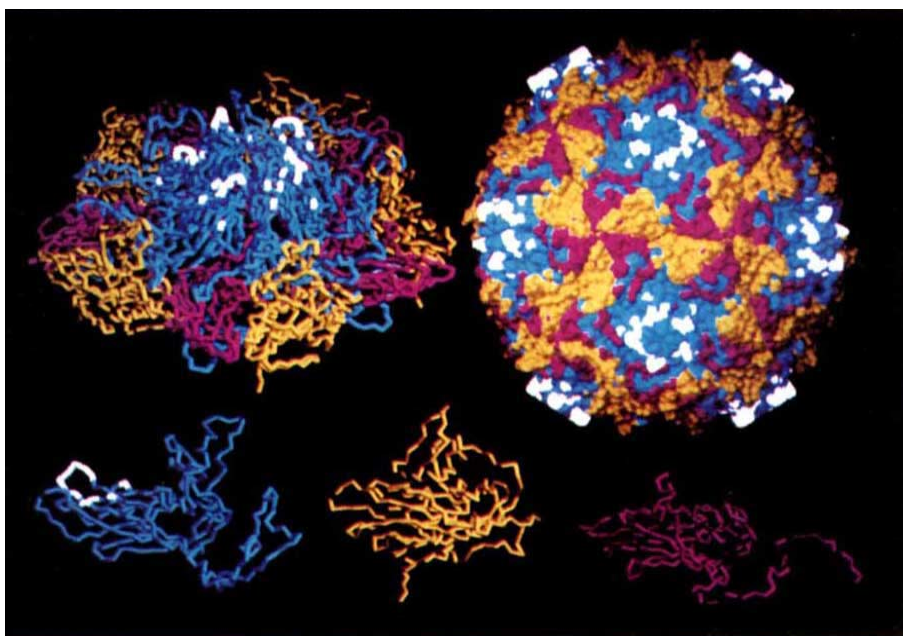
Poliovirus causes an estimated 250,000–2,000,000 cases of paralytic poliomyelitis annually, predominantly in unvaccinated or under-vaccinated children in developing countries. Introduction of both killed and live attenuated vaccines against poliovirus have resulted in the virtual elimination of poliomyelitis in industrialized countries. The live, attenuated polio vaccines are among the safest and most effective vaccines currently in use. Nonetheless, a handful of cases continue to occur each year in countries with extensive vaccination programmes. Virtually all

these cases are associated with an increased neurovirulence of the types 2 and 3 attenuated Sabin vaccine strains on replication in vaccinees³. In contrast, there is no well-documented vaccine-associated case of poliomyelitis caused by the Sabin type 1 strain³. Consistent with

this observation, genetic studies of the vaccine strains show that only two of the genomic differences between the Sabin 3 strain and its neurovirulent parent Leon 3 contribute significantly to its attenuation⁴, whereas the Sabin strain of type 1 poliovirus has accumulated multiple attenuating mutations with respect to its neurovirulent parent (type 1 Mahoney)⁵. It has therefore been suggested that an appropriate grafting of neutralizing antigenic sites from the genetically less stable type 2 and type 3 strains into the stable attenuated background of type 1 poliovirus could provide a vaccine with a considerably lowered incidence of reversion to neurovirulence.

In their paper in this issue², Burke *et al.* report the construction of an antigenic chimaera in which they replaced amino-acid residues 95–102 of VP1 of the type 1 Sabin strain with the corresponding sequence from a variant of the type 3 Sabin strain. The chimaera was constructed by oligonucleotide-directed mutagenesis, and has the expected reactivities with panels of antisera and with panels of monoclonal antibodies against type 1 and type 3 poliovirus. The chimaera produced titres (measured in an antigen-blocking assay) against both type 1 and type 3 poliovirus in rabbits and mice, and against both type 1 and type 3 poliovirus after oral administration to a cynomolgus monkey.

Murray, Kuhn and Wimmer (State University of New York at Stony Brook), collaborating with Kawamura and Nomoto (Tokyo Metropolitan Institute of Medical Science) and with Arita (Japanese



The region that has been replaced in the antigenic chimaeras described here is a 'loop' in the three-dimensional structure of the Mahoney strain of type 1 poliovirus⁶. Capsid proteins VP1, VP2 and VP3 are blue, yellow and red, respectively. The substituted loop in VP1 is highlighted in white. The internal protein VP4 is not shown. Alpha-carbon models of the individual capsid proteins are shown at the bottom; a pentamer is shown at the top left; an intact virion is shown at the top right. Note that five copies of the substituted loop cluster near each of the 5-fold axes of the particle.

*Molecular aspects of picornavirus infection and detection, ICN-UCI international conference on virology, Newport Beach, California, 14–15 January 1988.

National Institute of Health), used a different strategy to produce a similar chimaera. They constructed a mutagenesis cartridge using a naturally occurring and an artificially constructed restriction site in the infectious complementary DNA clone of the Mahoney strain of type 1 poliovirus which can accept synthetic double-stranded DNA fragments. Using this strategy, they produced a chimaeric virus in which residues 88–102 of VP1 of the Mahoney strain of type 1 poliovirus are replaced by the corresponding amino acids from the Sabin strain of type 3 poliovirus. This chimaera is neutralized by antisera against both type 1 and type 3 poliovirus, and induces a significant neutralizing response to both type 1 and type 3 poliovirus in rabbits and monkeys.

Girard, Martin, Couderc, Crainic and Wychowski (Pasteur Institute) used a similar mutagenesis cartridge, in this case with two artificially generated restriction sites, to replace residues 94–102 of VP1 of the Mahoney strain of type 1 poliovirus with the corresponding residues from the mouse-adapted Lansing strain of type 2 poliovirus. This chimaera is neutralized by both type 1 and type 2 antisera and generates a neutralizing response to both serotypes in rabbits. Moreover, although most strains of poliovirus (including the Mahoney strain of type 1 poliovirus) are specific for primates, the substitution of residues 94–102 from the Lansing strain of type 2 poliovirus confers on the chimaera the ability to cause paralytic poliomyelitis in mice. Earlier work by Racaniello and colleagues (Columbia University) shows that mouse-adaptation in the Lansing strain of type 2 poliovirus maps to the capsid region of the genome⁶, and that mutations in the 94–104 loop of VP1 significantly diminish the virulence of the Lansing strain in mice⁷. The mouse-virulence of the chimaera demonstrates conclusively that the 94–104 loop of VP1 is in itself an important determinant of the host range of poliovirus, and indicates that this highly exposed loop (see figure) may be involved in receptor binding, at least in the mouse central nervous system.

The success of these three groups in producing viable intertypic chimaeras of poliovirus demonstrates the feasibility of this approach for producing improved poliovirus vaccines. As a practical matter, however, the efficacy and safety of the existing poliovirus vaccines may make the

large-scale testing of any new polio vaccine difficult. Nevertheless, the results do allow considerable optimism that similar approaches may permit the easily grown and (relatively) genetically stable Sabin 1 strain to act as a carrier of antigenic determinants of other viruses, such as hepatitis A virus, which do not have vaccines and which are sufficiently difficult to grow to inhibit attempts to produce a vaccine by conventional methods. Finally, the exciting demonstration that the

Lansing–Mahoney chimaera is mouse-adapted is a poignant reminder that the usefulness of these chimaeras extends well beyond their immediate implications for vaccine development, as they are very accurate probes of the roles of specific amino-acid sequences in determining virus structure and biological functions. □

James M. Hogle is in the Department of Molecular Biology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, California 92037, USA.

Superconducting ceramics

Rare-earth elements redundant

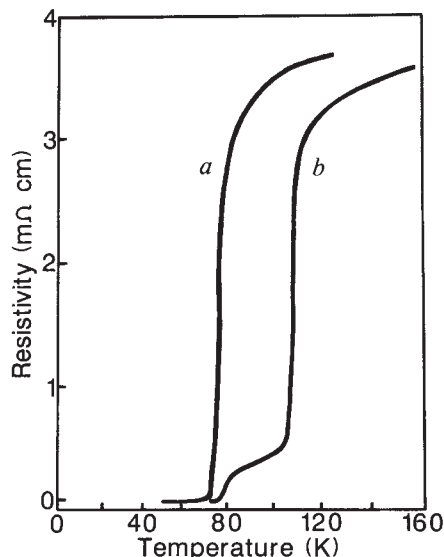
Ted Forgan and Colin Greaves

DESPITE the legions of researchers now studying high-temperature superconductors, there have been no reproducible measurements of superconductivity at temperatures above 93 K, the record established with $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ (YBCO) a year ago. This situation now seems set to change. On page 55 of this issue¹, Sheng and Hermann report that Tl-Ba-Cu-O becomes superconducting below 85 K, and their forthcoming results² indicate that partial substitution of calcium for barium increases the transition temperature (T_c) to around 115 K. Also, Maeda *et al.* can now observe³ superconductivity at 105 K in Bi-Sr-Ca-Cu-O . This increase in T_c above 93 K, although modest, is nevertheless important because it doubles the operating margin between T_c and 77 K, the boiling point of liquid nitrogen. Moreover, unlike the ceramics previously reported to have higher transition temperatures, the materials seem to be stable and, for the bismuth system at least, the results have been confirmed in many laboratories. In addition, the high- T_c phenomenon is no longer constrained to compounds of the rare-earth elements, so that a wider spectrum of possibilities for research now exists.

In a paper published⁴ at the end of 1987 (but submitted in May), Michel *et al.* reported superconductivity in Bi-Sr-Cu-O , although the transition temperatures were disappointingly low: 7–22 K. The apparent relationship between this system and the previously reported high- T_c mixed copper oxides stimulated Maeda and co-workers at Tsukuba in Japan to try other substitutions, which resulted in superconducting transitions near liquid-nitrogen temperatures in a Bi-Sr-Ca-Cu oxide sample prepared on 23 December 1987. This result was revealed to the Japanese press on 22 January 1988 and was confirmed within days in several Japanese laboratories. On 25 January, Chu and co-workers in Houston announced similar results, but with Bi-Sr-Ca-Cu-Al-O .

The work by Sheng and Hermann reported¹ in this issue follows a different path: they attempted to replace Y^{3+} in YBCO by another trivalent cation, and on 22 January, reported 85 K superconductivity in Tl-Ba-Cu oxides, and a few weeks later they announced superconducting behaviour at around 115 K in the similar Tl-Ba-Ca-Cu oxides. Once again, the importance of chemical intuition and knowledge of the periodic table in synthesizing new high- T_c materials is evident. Interestingly, however, attempts to substitute cations have often resulted in new phases rather than the intended doped versions of the parent phase.

It appears that the Bi-Sr-Ca-Cu-O samples are multiphase with two significant superconducting phases; the one with the lower T_c , around 80 K, is dominant in most samples synthesized so far and its complete characterization is probably imminent. Our X-ray diffraction patterns reveal a close structural relationship to the Aurivillius phases⁵, which were



Resistivity curves for $\text{BiSrCaCu}_2\text{O}_x$ showing both the single phase sample (a) with $T_c = 75$ K and a multiphase sample (b) with higher $T_c = 105$ K. (From ref. 3.)

1. Minor, P.D. *et al.* *J. gen. Virol.* **67**, 1283–1291 (1986).
2. Burke, K.L., Dunn, G., Ferguson, M., Minor, P.D. & Almond, J.W. *Nature* **332**, 81–82 (1988).
3. WHO Collaborative Study. *J. biol. Stand.* **9**, 163–184 (1981).
4. Westrop, G.D. *et al.* in *The Molecular Biology of Positive Strand Viruses* (eds Rowlands, D.J., Mahy, B.W.J. & Mayo, M.) (Academic, London, 1986).
5. Omata, T. *et al.* *J. Virol.* **58**, 348–358 (1986).
6. La Monica, N. *et al.* *J. Virol.* **57**, 515–525 (1986).
7. La Monica, N. *et al.* *Virology* **161**, 429–437 (1987).
8. Hogle, J.M., Chow, M. & Filman, D.J. *Science* **229**, 1358–1365 (1985).

previously used as a model for the Bi–Sr–Cu–O superconductor⁴. In these phases, $\text{Bi}_2\text{O}_{2.2+}$ layers separate up to four perovskite-like layers (three in the 80 K phase) in which copper ions occupy at least some of the small octahedral sites. Related Cu–O layers are responsible for the two-dimensional planes and one-dimensional chains which appear to be of fundamental importance in the previously discovered high- T_c superconductors. It is therefore fairly certain that similar structural features are present in the new materials. Electron-, X-ray- and neutron-diffraction studies should soon give a much clearer picture (W. I. F. David and D. McK. Paul, personal communication; R. M. Hazen *et al.*, preprint). The thallium-containing materials seem to have a different structure (A. M. Hermann, personal communication) and confirmation of structural data for this class is eagerly awaited.

A frustrating feature of the YBCO superconductor family is that almost every effective substitution results in a decrease in T_c : YBCO seems to represent a 'local peak' of T_c in the range of readily adjustable parameters. To discover new high- T_c families, it has been necessary to descend into a valley before finding a new area of high ground. We are now experiencing a return of the old excitement as researchers cast around to see if there are new, undiscovered peaks. In this respect, note that the new bismuth-containing materials seem to be very tolerant of variations in composition, and the thallium-based materials also seem to belong to an extended family.

One aspect of the new broad field of investigation is the vital need for optimization of preparative conditions. Lower temperatures are certainly required than for YBCO, and this in itself is an attractive feature. Also of potential significance is the apparent stability of the new materials, especially the resistance of the bismuth-based phases to degradation by water. The importance of the recent discoveries is further underlined by the fact that the increase in T_c could be very helpful in enhancing the critical currents so far achieved with ceramic superconductors. Finally, because the new T_c is above the boiling point of liquid methane (110 K), one new application immediately suggests itself: the measurement of resistance as a level detector in tanks of liquefied natural gas. □

1. Sheng, Z. Z. & Hermann, A. M. *Nature* **332**, 55–58 (1988).
2. Sheng, Z. Z. & Hermann, A. M. *Nature* (in the press).
3. Maeda, H., Tanaka, Y., Fukutomi, M. & Asano, T. *Jap. J. appl. Phys. Lett.* (in the press).
4. Michel, C. *et al.* *Z. Phys.* **B68**, 421 (1987).
5. Aurivillius, B. *Arkiv. Kemi.* **1**, 463 (1949).

Ted Forgan is in the Department of Physics and Colin Greaves is in the Department of Chemistry at the University of Birmingham, PO Box 363, Birmingham B15 2TT, UK.

Biological optics

Paradoxical superposition

Michael F. Land

THE discovery of a new type of imaging system in an eye is always an event worth celebrating. By my count there were only about eight such systems until today, and now there are nine, thanks to D.-E. Nilsson's careful study of the eye of the crab *Macropipus* on page 76 of this issue. I cannot claim to like this eye, however. It is quite the most complicated optical system known in biology, and resolutely

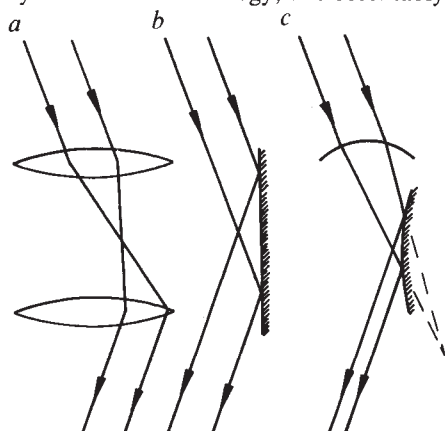


Fig. 1 Three ways of inverting the direction of a light beam, corresponding to the three types of superposition eye. a, Refracting; b, reflecting; c, parabolic.

difficult to understand. It was even hard to find a name for it, and although "parabolic superposition" finally won over "peculiar superposition" and "problematic superposition", those of us who watched, impressed, as Nilsson struggled with this eye felt that so straightforward and informative a name failed to do justice to its real deviousness. The feature of the eye that makes it particularly alarming for biologists, although intriguing for optical technologists, is that the mechanism of bending light is quite different in different planes.

These eyes are a variant of the superposition type of compound eye, in which many facets contribute ray-bundles to a single erect image on the deep-lying retina. The essential feature of the optical components of these eyes is that they redirect light-beams across their axes, and the way they do this defines the type of eye (Fig. 1). Until today two kinds of superposition eye have been recognized. The first, found in moths and krill, relies on lenses arranged in pairs as an array of tiny inverting telescopes (Land, M. F. *Nature* **287**, 681–686; 1980). The second, found in shrimp and crayfish, uses a square array of plane mirrors to achieve the same inversion. The mirror type was actually the last natural optical system to be discovered, by Klaus Vogt (*Z. Naturforsch.* **30c**, 691; 1985). I

remember wondering at that time whether intermediates were possible — whether some combination of a single lens and half a mirror could be persuaded to do the same job — but I abandoned the speculation because nothing I could think of corresponded to the half-mirror. Confident the laws of physics were on my side, I concluded that intermediates could not exist.

The upshot of Nilsson's study is that my confidence was misplaced. There is at least one, and possibly many solutions to the problem. In the version reported here, the crab *Macropipus* makes use of a single lens and of a mirror surface with a unique form, cylindrical in cross-section but a convex parabola in profile (Fig. 2). The parabola straightens out the converging beam from the lens (Fig. 1c), so that the combination behaves much like a plane mirror. The tricky bit comes when the beam is viewed from along the axis of the element. The mirror surface is now not parabolic and convex but circular and concave, and the body of the element itself behaves as a cylindrical lens. The result is an optical system that as before produces a parallel emergent beam, but using a lens–mirror–lens combination, not a lens and parabolic mirror.

Why should some crabs use this complicated mechanism? The answer seems to be that it gives them two alternative types of eye. Under bright conditions, dark pigment intercepts the reflected rays, leaving only those that are focused by the corneal lens onto the rear surface of each element. Here a light-guiding structure conveys the focused light down to the

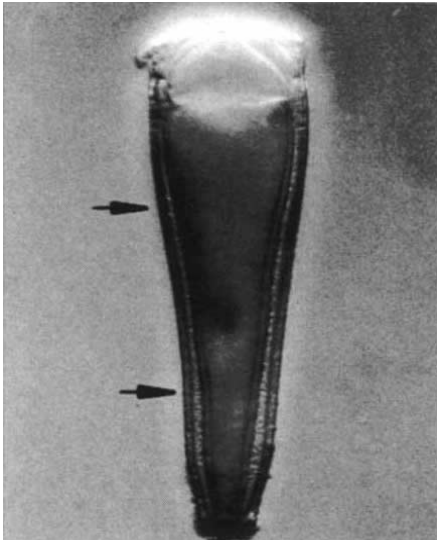


Fig. 2 An optical element (crystalline cone) from the crab *Macropipus* showing the region of parabolic curvature (arrows). (Photograph by D.-E. Nilsson.)

rhabdom. In this configuration the eye acts as an ordinary apposition eye (like that of a bee, for example) with each rhabdom having its private optical system. When the screening pigment retracts in the dark the reflected, superposition rays take over, forming a much brighter image directly on the rhabdom layer. From a taxonomic point of view these eyes also represent an interesting compromise, being a clear intermediate between the reflecting superposition eyes of the long-bodied decapods (shrimps and crayfish) and the other crabs with pure apposition eyes. In some ways this is sad. A short while ago we could say that "X could not evolve into Y because the eye of the intermediate

form would not work". Now we cannot; Nilsson's eye links apposition eyes to both other types of superposition eye, acting as a kind of universal taxonomic glue.

It is too early to tell whether the parabolic superposition principle will be of value in optical technology. I have failed to find anything like it in optics textbooks. Its cousins have been useful: refracting superposition mechanisms have influenced inhomogeneous lens design, and the reflecting mechanisms have found a use in X-ray telescopes. The trick is to find a problem to which this eye is the solution. □

Michael F. Land is Professor of Biology at the University of Sussex, Falmer, Brighton BN1 9QG, UK.

Agriculture

Models of crop growth

Gareth Hughes

THE first step in appraising the effect of a pest or disease on crop yield is usually to characterize the relationship between severity of attack and loss in yield¹. Most people have used an empirical regression equation to couple crop yield to pest and disease dynamics, even in cases where relatively realistic models of the latter are available. Recently, however, interest has grown in developing simulations of crop growth driven by plant physiological processes in relation to the physical environment, and which include mechanistic models of pests and diseases to enable the assessment of losses². But such models can often be too demanding in terms of input data requirements to be useful in practice³. Waggoner and Berger now suggest⁴ that the effects of pests and diseases can be modelled in a similar way to the effects of agronomic factors such as plant population density and nutrient and water supply on crop growth and yield.

The paradigm for crop-growth models is Monteith's analysis⁵ of biomass (W) as the time-integrated product of the photosynthetically active radiation (PAR) incident on the top of the leaf canopy (R), the fraction (f) of PAR intercepted by the canopy and the amount of dry matter produced per unit PAR intercepted (e). Many analyses have shown that, at least under good conditions, W is almost proportional to the amount of intercepted PAR ($I = \int f.R.dt$), where e is the proportionality constant and t time⁶. In such cases, crop yields depend on the growth and duration of the leaf canopy in relation to the seasonal pattern of irradiance. Recent data show⁷ that physical environmental stresses can reduce crop yields by their effects on both I and e . The question raised⁴ by Waggoner and Berger concerns the nature of the effects of biotic stresses on these variables. They reanalysed data

from four crops — peanut, potato, maize and wheat — under pathogen attack, with the objective of characterizing relationships between W and the amount of radiation intercepted by healthy foliage. These relationships all seemed to be linear, and no differences were apparent for cases where data were also available for a comparison of the slopes (estimates of e) of relationships for diseased and manually defoliated crops. Haverkort and Bica-mumpaka reached a similar conclusion⁸



The aftermath of pest disturbance of an experimental crop of peas at Finca Zamadueñas, Spain.

from experiments in which they grew potato crops with or without control measures against *Phytophthora infestans*. They could account for practically all the differences in yield by differences in radiation interception by healthy leaves. Both diseased and healthy crops had the same value of e .

Two aspects of these results^{4,8} are particularly noteworthy. First, bearing in mind differences in photosynthesis that have been detected in healthy and diseased plants⁹, it is interesting that no differences in e between healthy and diseased crops were detected. This is good news for those

who develop crop-loss assessment models using manual defoliation to simulate the effects of pests and diseases, although examples of variation in e in attacked crops will surely emerge.

The second interesting feature of the new results is the apparent insensitivity of the underlying analysis of crop growth to changes in the spatial pattern of the crop. Most pests and diseases have patchy spatial patterns, and thus can disrupt the more-or-less regular patterns in which crops are sown, in addition to inflicting plant injury (see figure). This is recognized as a problem in both conventional¹ and crop-growth-based² loss-assessment modelling, because the spatial pattern of plants can influence the level of intra-crop competition. Some recent results¹⁰ show how patchiness in a crop can influence yield, independent of plant population density (and pests and diseases).

Waggoner and Berger's analysis⁴ suggests that the function relating f to green-area index in healthy crops can also serve for diseased crops. They do not account for any differences in radiation interception resulting from the spatial patterns of the foliage of healthy and diseased crops. Haverkort and Bica-mumpaka⁸ estimate f directly, rather than as a function of green-area index, but this cannot account for changes in the level of intra-crop competition for resources other than PAR resulting from patchiness in attacked crops.

Neither analysis includes a crop response to plant spatial pattern, and results from studies¹¹ in which the effects of spatial pattern as well as severity of crop injury are incorporated into conventional loss-assessment models demonstrate the importance of including both. There are new models relating crop growth and yield to plant spatial arrangement and resource capture^{12,13} which, although they do not deal directly with the problem of patchiness arising from pest or disease disturbance, may provide a framework for the development of crop-growth-based loss-assessment models that include responses to the spatial pattern of crop injury. □

1. James, W.C. *A Rev. Phytopath* **12**, 27-48 (1974).
2. Boote, K.J. *et al. Phytopathology* **73**, 1581 (1983).
3. Teng, P.S. in *Crop Loss Assessment and Pest Management* (ed. Teng, P.S.) 105-113 (Am. Phytopath. Soc., 1987).
4. Waggoner, P.E. & Berger, R.D. *Phytopathology* **77**, 393-398 (1987).
5. Monteith, J.L. *Phil. Trans. R. Soc. B* **281**, 277-294 (1977).
6. Monteith, J.L. & Elston, J. in *The Growth and Functioning of Leaves* (eds Dale, J.E. & Milthorpe, F.L.) 499-517 (Cambridge University Press, 1983).
7. Russell, G., Jarvis, P.G. & Monteith, J.L. in *Plant Canopies, Their Growth Form and Function* (eds Russell, G. *et al.*) (Cambridge University Press, in the press).
8. Haverkort, A.J. & Bica-mumpaka, M. *Neth. J. Pl. Path.* **92**, 239-247 (1986).
9. Buchanan, B.B. *et al. in Effects of Disease on the Physiology of the Growing Plant* (ed. Ayres, P.G.) 13-28 (Cambridge University Press, 1981).
10. Soman, P. *et al. Agron. J.* **79**, 891-895 (1987).
11. Campbell, C. & Noe, J.A. *Rev. Phytopath.* **23**, 129 (1985).
12. Benjamin, L.R. & Hardwick, R.C. *Ann. Bot.* **58**, 757 (1986).
13. Sutherland, R. & Benjamin, L. *Ann. Bot.* **59**, 399 (1987).

Gareth Hughes is at the Edinburgh School of Agriculture, University of Edinburgh, West Mains Road, Edinburgh EH9 3JG, UK.

Earthquakes

Seismic potential of the Cascadia subduction zone

Garry C. Rogers

THE two main conclusions that emerged from a special session at a recent meeting* are that the Cascadia subduction zone appears to be storing strain, and that there are numerous examples of suddenly buried wetland-soil horizons that seem to be best explained as examples of past great earthquakes. Although some speakers suggested that there are special conditions that could prevent great earthquakes, such as the high temperatures of the young subducting slab (J. Seeveringhaus, University of California, Santa Barbara) or the abundant offshore sediments (D.E. Byrne, Lamont-Doherty Geological Observatory; L.R. Sykes, Lamont-Doherty), none offered alternative explanations for the data that support the 'great earthquake' hypothesis. The outcome of the debate has more than passing interest for the ten million people who live in the vicinity of the zone and could be affected by a great earthquake.

The Cascadia subduction zone is a small subduction zone that straddles the Canada-United States border on the west coast of North America (see figure). It involves several small oceanic lithospheric plates (the last remnants of the large Farallon plate) that are actively subducting beneath North America. Low-angle thrust earthquakes, characteristic of most subduction zones around the world, have not been observed there. Detailed monitoring by seismic networks on both sides of the border over the past decade shows that, although small earthquakes are common in the overlying and descending plates, thrust earthquakes at the subduction interface are absent even at the microearthquake level. The question thus arises as to whether subduction is taking place aseismically or whether the region is in an interseismic period between great earthquakes. With this seismological information, strain measurements and studies of the recent (the past 12,000 years) geology of the region becoming available, the question could soon be answered.

The search for geological evidence of past great earthquakes is focused on identifying coastal features that indicate sudden changes of elevation relative to sea level. There are no recent marine terraces along the Cascadia subduction zone to suggest that sudden emergence events have occurred (D.O. West, Golder Associates). But the hypothesis put

forward by B.F. Atwater, of the US Geological Survey (USGS), in *Science* (236, 942-944; 1987), that sequences of buried wetland soils on the outer coast of Washington State represent a series of sudden down-drops generated by large subduction earthquakes, received considerable support.

A search for such buried-soil horizons by several investigators reveals that they extend from southern Oregon (A.R.



The Cascadia subduction zone of western North America. Oceanic lithosphere created at the nearby Juan de Fuca Ridge system subducts beneath the North American Plate.

Nelson, USGS) to southern Vancouver Island (my own work). The most complete sequence has been found in large estuaries on the outer coast of Washington (Atwater; A.G. Hull, University of California, Santa Barbara) and indicates that submergence events occurred about 300, 1,000, 1,500, 1,700, 2,500, 2,800 and 3,500 years ago. The presence of sand layers, interpreted by Atwater as tsunami deposits, help to mark the areal extent of some events (M.A. Reinhart, University of Washington; W.C. Grant, USGS). Investigations in smaller estuaries, particularly those in Oregon, are confusing, as buried soil layers are found, but not always (Nelson; Grant; C. Peterson, Oregon State University). It is clear that

Erratum

In the table of the article "How to live with radon" by Michael O'Riordan (*Nature* 331, 302; 1988), the term 'per cent' should have referred only to the absolute lifetime risk and not to the relative lifetime risk of exposure to radon. □

mapping the extent of these suddenly buried soils, dating them and further research into their origin is important.

The essential results of horizontal strain measurements in the subduction region are that wherever measurements have been attempted, strain appears to be accumulating with a principal axis of shortening that is parallel to the modelled direction of plate motion (M. Lisowski, USGS; H. Dragert, Geological Survey of Canada; N.A. Breen, Lamont-Doherty). Repeated horizontal control surveys comparing older surveys with modern laser ranging or global-positioning satellite measurements show that shear strain is accumulating at a rate of about 0.2 parts per million per year, comparable to rates in other subduction zones such as Japan, New Zealand and Alaska, where large thrust earthquakes have occurred in historic times. These measurements support the hypothesis that the subduction zone is locked and strain is currently accumulating in the upper plate.

Measurements of vertical motion are also consistent with strain accumulation across a locked subduction zone. Changes of sea level at tidal stations in the northern part of the zone reveal a characteristic two-dimensional pattern that would result from the oceanic plate dragging down the edge of the continental plate (my own work). Integrating precise levelling data with tide gauge data shows a complex areal distribution of vertical motion but one that is also consistent with strain accumulation (S.R. Holdahl, US National Geodetic Survey). The patterns of vertical and horizontal motion at the north end of the subduction zone are complicated by deformation caused by large crustal earthquakes (Holdahl; Dragert).

A two-dimensional finite element elastic model of the subduction zone with reasonable estimates for input parameters is found to fit observed strain measurements best when the subduction zone model is locked and close to an impending earthquake (H.J. Melosh, University of Arizona). Current observations, however, are too sparse to provide diagnostic constraints for the model.

Before the arguments presented at the meeting lead us to believe in a cycle of great earthquakes on the Cascadia zone, the sudden subsidence events must be dated accurately and correlated over a large area. Similarly, strain must be shown to be accumulating over a large area. The evidence seems to be sufficient, however, to place the onus on the sceptics of the great-earthquake hypothesis to provide convincing alternative explanations for the observations. Until now, none has emerged. □

Garry C. Rogers is at the Geological Survey of Canada, Pacific Geoscience Centre, Sidney, British Columbia V8L 4B2, Canada.

* American Geophysical Union fall meeting, 7-11 December 1987, San Francisco.

Immunology

Tolerance, restriction and the Mls enigma

Miranda Robertson

Two papers on pages 35 and 40 of this issue of *Nature*^{1,2} report strong evidence that immunological tolerance is due to the elimination of lymphocytes bearing receptors capable of recognizing self antigens. At first sight this is unremarkable. Three possible mechanisms have been proposed to explain tolerance: the inactivation of self-reactive lymphocytes (clonal anergy); their chronic suppression (clonal suppression); and their elimination (clonal deletion). Clonal deletion, as the simplest and safest of the three, is widely thought to be the most likely, and moreover there is already good evidence for its operation in the case of tolerance to the products of the self major histocompatibility complex (MHC)³. The extension of this evidence to a non-MHC antigen, which is the subject of the two reports in this issue^{1,2}, is plainly important but not in itself startling. What makes these papers particularly interesting is the light they may obliquely shed on the central paradox in the ontogeny of thymus-derived lymphocytes: namely, the need to select a repertoire of T cells that will recognize MHC molecules when they are associated with foreign antigens but ignore them when they are not. The need for such a repertoire arises from the special part played by MHC molecules in directing T-cell responses; and the implications of the recent experiments on tolerance arise from the nature of the antigen involved, which seems to violate the restrictions that normally govern MHC-directed immune responses.

Restriction and tolerance

All the functions of T lymphocytes depend upon interactions with other cells: cytotoxic T lymphocytes kill virus-infected cells, and helper T lymphocytes induce the proliferation of antibody-secreting B lymphocytes. Thus the activities of T lymphocytes must be focused on the surfaces of other cells; and it is the function of the MHC molecules to direct T-cell responses to the cell surface, where foreign antigens are recognized as part of a complex with the products of the MHC. Co-recognition by T cells of the MHC-antigen complex is normally both restricted to the MHC molecules expressed by a given individual, and highly specific. Because they must bind a wide range of different antigens, the several surface glycoproteins encoded by the MHC of any individual are extremely polymorphic, and the total universe of MHC-antigen complexes is correspondingly large.

Nonetheless, each T lymphocyte will normally be able not only to distinguish the presence or absence of foreign antigen on the MHC molecules of that individual (the prerequisite for self tolerance), but to discriminate exquisitely between specific combinations of antigen and MHC molecule.

It is the ability of the Mls antigens of some strains of mice to override this exquisite specificity that has made it possible to detect the elimination of T lymphocytes responsive to them; and has suggested a mechanism for reconciling the requirements for MHC restriction and self tolerance in thymic ontogeny.

The Mls enigma

Mls is the name given to a mouse antigen (or antigens) discovered by Hilliard Festerstein in the 1960s through its ability to stimulate MHC-identical lymphocytes to proliferate *in vitro*⁴. This ability is controlled by a single locus (called Mls for mixed lymphocyte stimulatory); and specifically, cells from mice carrying the Mls^a allele elicit vigorous proliferative responses from MHC-identical cells carrying the Mls^b allele. (For a clear summary of the prevailing picture of Mls, which is more complicated than I have acknowledged, see ref. 5.) These responses are even greater in magnitude than the allogeneic reactions between MHC-non-identical cells, and for this reason it has been assumed that, like alloreactivity, they must reflect a functionally important bias in the T-cell repertoire.

The magnitude of the response to Mls is due at least in part to its failure to respect the normal constraints of MHC restriction: mouse cells of a single MHC type will react to Mls^a not only on cells of that type, but also on cells of a range of other MHC types. The nature of the Mls protein is unknown; but because of this failure of restriction, one suggestion has been that Mls is recognized independently of the MHC-restricted antigen receptor, by a separate receptor subserving an accessory recognition function on the T cell⁶. The absence of any human equivalent to Mls, however, together with clear evidence for Mls-MHC co-recognition⁷, have militated against the acceptance of this idea; and now Kappler *et al.*¹ and MacDonald *et al.*² provide powerful evidence for the recognition of Mls by receptors belonging to the MHC-restricted T-cell repertoire, and thus that the failure of MHC restriction must be due to a special property of the

Mls-receptor interaction. Their evidence comes from an analysis of the Mls reactivity of T cells bearing specific identified receptor chains.

Clonal deletion

Both groups were led to their experiments by the adventitious observation that cells expressing specific T-cell receptor chains are absent from mice bearing Mls. The antigen receptor of T lymphocytes is a heterodimer of two chains, α and β , each with a variable region analogous to the variable region of immunoglobulin chains, the two variable regions together being presumed to form the antigen-binding site. Individual variable regions can be identified by monoclonal antibodies, and in the course of screening with such antibodies, Kappler *et al.* and MacDonald *et al.* found that mice expressing Mls^a have few or no T cells bearing receptors containing the specific β -chain variable regions V β 8.1 and V β 6.

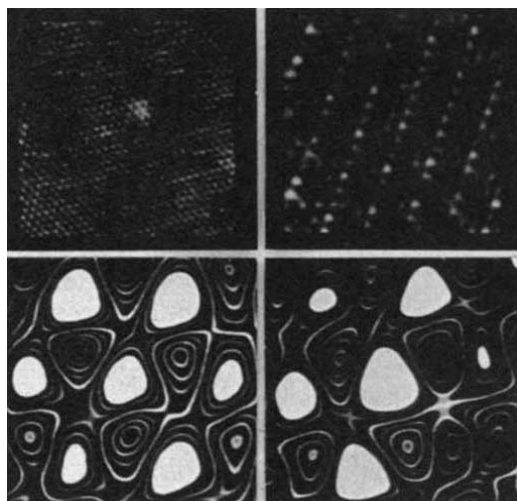
This suggested an analogy with the earlier observations of Kappler *et al.*³ on tolerance to the mouse MHC molecule I-E. The V β region V β 17, which confers on almost all receptors containing it the ability to bind to all alleles of I-E, is eliminated from the mature T-cell repertoire of mice carrying I-E alleles, despite the presence of the V β 17 gene segment in the genome of the mice. Because the product of the V β 17 segment can also be detected on immature thymocytes in such mice, Kappler *et al.* concluded that T cells bearing this V region are eliminated during thymic ontogeny in mice expressing I-E.

From this precedent, Kappler *et al.* and MacDonald *et al.* reasoned that the absence of V β 8.1 and V β 6 in the mature T-cell repertoire of mice expressing Mls^a might imply that these two V regions confer Mls^a reactivity on those cells bearing them; a supposition they were able to confirm by an analysis of hybridomas with receptors containing these V β regions — a very high proportion of which did indeed recognize Mls^a. In the case of V β 8.1, moreover, this response was shown to have the broad MHC specificity typical of Mls reactions.

Together, these data clearly show that a remarkably high proportion of mature T cells must bear $\alpha\beta$ receptors that interact so strongly with Mls^a as to override, at least in part, the characteristic specificity of the receptor-MHC interaction. These conclusions have inspired Kappler *et al.* to risk an interesting and far-reaching speculation on the structural basis of the mechanism whereby MHC-restricted antigen recognition is imposed on the T-cell repertoire during maturation in the thymus. The beginning of the chain of inference by which they arrive at their speculation lies in the nature of the interactions between the $\alpha\beta$ heterodimer on

Images of charge density waves

THE uses of scanning tunneling microscopes (STMs) are many: both the manipulation of single molecules and the study of surface topography have been discussed recently in News and Views. R. V. Coleman and co-workers now report (*Phys. Rev. B* 37, 2741; 1988) the use of one to study charge density waves (CDWs) on the surface of a metal cooled to 4.2 K. The periodic distribution of conduction electrons is determined primarily by the ions in the metal; but below a critical temperature, interactions mediated by lattice vibrations (phonons) between the electrons (the same interactions lie at the heart of superconductivity) can give rise to additional charge modulation — the charge density waves. Coleman and co-workers report the observation of CDWs in NbSe, which consists of layers of niobium sandwiched between layers of selenium ions. At low magnification (top left, $94 \times 94 \text{ \AA}^2$), the STM reveals the lattice of selenium ions at the surface, observed by the increased tunnelling current detected by the scanning tip as it passes over the ions. (In practice, the tip height is varied to maintain a constant current.) The CDWs, however, become evident only at higher magnification (top right, $55 \times 55 \text{ \AA}^2$): there is an additional modulation of the charge density with a period of three times the lattice spacing. The lower micrographs show the charge distribution at the crest (right) and minimum (left) of the CDWs; the central selenium atom on the crest appears to be 'higher' than the surrounding atoms. The distortion of the charge distribution ('leaning' to the right) arises because the CDWs are centred on the offset niobium atoms in the second layer of the crystal and not on the surface selenium ions.



R. V. Coleman

the responding lymphocyte, and the Mls^a-MHC complex on the target cells.

Anatomy of recognition

The inferential nature of the scheme is imposed partly by the absence of direct structural information about the MHC molecules that engage in Mls reactions. There are two classes of MHC recognition responses, broadly corresponding to the two major functional classes of T lymphocyte. Cytotoxic cells generally recognize antigens bound to MHC class I molecules, while helper T cells generally recognize antigens bound to class II MHC molecules. In both cases it is known that the antigen bound to the MHC molecule is a short peptide derived by intracellular degradation from the complete protein antigen, but only for the class I molecule has the crystal structure been solved^{8,9}. Because the responses stimulated by Mls^a are helper-cell responses directed against class II molecules, it has been necessary to infer the structural basis of antigen presentation in Mls^a reactions on the basis of structural homologies between class I and class II that suggest that they bind antigen in the same way.

The crystal structure of the class I molecule has revealed that the antigenic peptide is bound in a pocket on the external surface of the molecule, the sides

of the pocket being formed by polymorphic residues implicated in interactions with the T-cell receptor^{8,9}. Thus the complex of MHC and antigen presented to the T cell consists of one surface of the antigenic peptide flanked by polymorphic residues of the MHC molecule.

Thus, as receptor binding normally depends upon the sum of the contacts made by the MHC molecule and the peptide antigen, the implication is that Mls binds so strongly to the antigen-binding residues of the T-cell receptor that the effect of specific contacts between receptor and MHC is negligible. Hence the violation of MHC restriction. The final link in the chain of inference leads to the question of how this may bear on the expansion of the self-restricted T-cell repertoire in the thymus.

Thymic maturation

According to a current model of T-lymphocyte ontogeny (lucidly summarized in ref. 10), maturation in the thymus occurs in two stages. In the first, functionally immature thymocytes expressing $\alpha\beta$ receptors are positively selected on thymic epithelial cells for weak binding to self MHC molecules, giving rise to self-MHC restriction; in the second, they are negatively selected in the thymic medulla for strong binding to self-MHC molecules,

so that potentially self-reactive cells are eliminated and self tolerance is assured.

The nature of the selective mechanism is unknown: it is assumed that thymic epithelial cells induce proliferation in immature lymphocytes binding weakly to MHC molecules on their surface, while bone-marrow derived cells in the medulla kill those cells that bind strongly to their MHC molecules. But whereas it is easy to imagine a mechanism for reliably selecting strongly binding cells, it is more difficult to imagine how cells could be selected for a binding affinity so weak as to be ineffectual in the absence of antigen.

It is this difficulty that Kappler *et al.* attempt to address by invoking the Mls phenomenon. If the antigen-binding pocket of MHC molecules on thymic epithelial cells is normally occupied by a peptide or peptides with an affinity for the T-cell receptor strong enough to ensure binding in the presence of weak receptor-MHC interactions, then the consequence would be positive selection for weak binding to self MHC.

That the antigen-binding pocket of thymic epithelial cell MHC molecules should be permanently occupied by peptides specific to thymic epithelium is not as gratuitous a supposition as it may seem. There is evidence of various kinds suggesting that the antigen-binding pockets of MHC molecules are unlikely ever to be empty, and thus may normally be occupied by peptides derived from self proteins: the human class I molecule crystallized by Bjorkman *et al.*^{8,9} turned out to contain bound peptide, or peptides, despite extensive purification.

What Kappler *et al.* propose is that Mls^a has properties that adventitiously mimic the properties of a thymic epithelial cell protein that has evolved to bind with high affinity both to MHC and to the T-cell receptor antigen-binding site, the latter perhaps through conserved residues in this otherwise hypervariable region.

Under this supposition, Mls^a in mice might represent either the aberrant expression of such a thymic peptide in non-thymic cells; or a mutation in some normal cell protein that confers upon it the properties of such a peptide. In any case, it is plain that isolation of the Mls product and analysis of antigenic peptides derived from it is likely to become a significant goal of molecular immunology. □

1. Kappler, J. W., Staerz, U., White, J. & Marrack, P. *Nature* 332, 35–39 (1988).
2. Macdonald, H. R. *et al. Nature* 332, 40–45 (1988).
3. Kappler, J. W., Roehm, N. & Marrack, P. *Cell* 49, 273–280 (1987).
4. Festenstein, H. *Transplant. Rev.* 15, 62–88 (1973).
5. Janeway, C. *et al. Immun. Today* (in the press).
6. Katz, M. E. & Janeway, C. A. *J. Immun.* 134, 2064–2070 (1985).
7. Jones, B. & Janeway, C. A. *Immunogenetics* 16, 243–255 (1982).
8. Bjorkman, P. *et al. Nature* 329, 506–512 (1987).
9. Bjorkman, P. *et al. Nature* 329, 512–518 (1987).
10. Marrack, P. & Kappler, J. *Science* 238, 1073–1079 (1987).

Miranda Robertson is Biology Editor of Nature.

Edmund Brisco Ford (1902–1988)

EDMUND BRISCO ("Henry") Ford FRS died in Oxford, where he had worked all his life, on 22 January, aged 86. He had been Professor of Ecological Genetics at the University and was a Fellow of All Souls College, where he had served two terms as Senior Dean.

He had two outstanding achievements. First, he founded ecological genetics, demonstrating that the brilliant statistical methods and theoretical predictions of Fisher and Haldane could be used and tested in nature as well as in the laboratory, and the Lepidoptera constituted his main experimental material. Second was his conception of balanced polymorphism applied to the study of evolution in the wild. On both topics, which were inter-related, he wrote with marvellous clarity and was the supreme example of the benefits to a scientist of a classical education.

As a boy, Henry and his father, H.D. Ford, observed each season a colony of the Marsh Fritillary butterfly in Cumberland in the north of England. The numbers fluctuated greatly and during periods of rapid increase there was an extraordinary outburst of variability in pattern. When the population decreased again the common form was recognizably distinct from that which had prevailed before the period of abundance. An opportunity for evolution had occurred and the insect had made use of it.

A similar principle was used later in detailed studies on the frequency of the heterozygote *medionigra* of the Scarlet Tiger moth, *Panaxia dominula*, and on spot number in the Meadow Brown butterfly, *Maniola jurtina*. By this time Henry had devised elaborate techniques of mark-release-recapture, which enabled his team to estimate changes in frequency of particular forms, and of the genes controlling them, and to assess migration.

Ford was the first to predict that the polymorphic human blood-group systems

would influence susceptibility to disease. The associations of cancer of the stomach and group A, and of duodenal ulcer and group O, bore this out. In the sickle-cell haemoglobinopathy the advantage of the heterozygote was excellently demonstrated as it protected children against malaria.

In his later years Ford became interested in the genetics of the Gypsy moth, *Lymantria dispar*. Using the heterozygotic body technique, he and C.A. Clarke showed that Goldschmidt was wrong in thinking that unusual ratios in race crosses of the moth were the result of complete sex reversal. In fact the all-male broods were fully fertile and simply the result of the Haldane effect.

Henry had a gift for picking good research workers and then giving them their heads. H.B.D. Kettlewell thrived on the Peppered moth and industrial melanism; Arthur Cain excelled on the snail *Cepaea nemoralis*, demonstrating that natural selection by predators acted on a colour polymorphism; and Philip Sheppard concentrated on the evolution of mimicry, particularly in the Swallowtail butterfly *Papilio dardanus*, on which Ford's monograph was the Bible. But there was more to come: Sheppard applied Ford's suggestion about the human blood groups to the Rh (rhesus) system, and with researchers at the Department of Medicine at Liverpool University found a method of preventing Rh haemolytic disease of the newborn. It was for this type of work that the Nuffield Foundation, of which Henry was a trustee, set up the Unit of Medical Genetics there, and it was a nice quirk that in the Rh polymorphism the heterozygote does not obey the rules, for it is always at a disadvantage.

Henry's interests were very wide; he was an expert on archaeology and heraldry and with J.S. Haywood he wrote *Church Treasures in the Oxford District*. On genetics and ecology his many books are classics: *Mendelism and Evolution*; *The Study of Heredity*; *Genetics for Medical Students*; *Butterflies*; *Moths*; *Ecological Genetics*; *Genetic Polymorphism*; *Genetics and Adaptation*; *Understanding Genetics and Taking Genetics into the Countryside*. *Butterflies*, to his astonishment, was a best-seller.

Oxford has produced many eccentrics, but few in the same class as Henry. Though a very faithful and generous friend of people he liked, he froze those of whom he disapproved, and his attitude to women (he was unmarried) was well exemplified at one of his lectures when only female students turned up: "Since there is no one here, there will be no lecture". We shall never see his like again. **Cyril Clarke**

Sir Cyril Clarke is Director of the Research Unit at the Royal College of Physicians, 11 St Andrew's Place, Regent's Park, London NW1 4LE, UK.

Daedalus

Vehicular spectroscopy

MOTOR vehicles are notoriously unreliable. They all need regular servicing, and some countries impose an annual roadworthiness test for every vehicle. So Daedalus is seeking improved diagnostic methods. Every solid object has a huge set of mechanical resonances, amounting to a detailed 'signature' of its total structure. DREADCO's engineers are coupling an experimental car to an electromechanical frequency-generator, and sweeping this through a wide range of frequencies from sub-audio to ultrasonic. The vibrational amplitude and phase excited at many points of the vehicle are recorded by piezo-electric transducers. The whole test takes only a few minutes.

The resulting mechanical spectrum is then compared by computer to the one established for a perfect car of that make and model. Mass-production is so consistent that even small deviations from the standard will indicate some sort of anomaly, though detailed knowledge will be needed to distinguish between some harmless dimensional variation and a dangerous deficiency. Each part of the spectrum will tell its own story. The lower frequencies will be dominated by whole-body torsions and panel resonances, thus revealing the condition of the bodywork. A rusty area at a vibrational antinode, for example, would raise the relevant frequency; at a node it would lower it. The higher frequencies will give information about the more massive and rigid components, for example, the engine block and drive train. The amplitudes of some frequencies will be sensitive to mechanical damping, as contributed by shock-absorbers, underseal and so on. Overtones at any frequency will crucially reveal non-linearities, caused by the rattling of loose bolts, the opening and closing of tiny cracks, or the slapping of worn bearings. Comparisons with spectra recorded at previous services would show up such deteriorations even more sharply.

Thus all exploratory grubbing-about will be mercifully ended, and mechanics will spend all their time rectifying faults that are clearly indicated. So accurate is frequency-measurement that nothing can be overlooked; once the spectrum agrees with specification, the proud owner will be able to drive away in a 'spectroscopically pure' car. But, says Daedalus, modern instrumental methods permit still greater elegance. Continuous-wave spectroscopy could with advantage be replaced by fast-Fourier techniques. In this method of instant diagnosis, the wired-up vehicle will simply be given a single heavy blow, and the full vibrational spectrum will be reconstructed from the dying echoes by Fourier transformation. **David Jones**



100 years ago

MR. SHELFORD BIDWELL is continuing his admirable researches on the changes produced by magnetism in the lineal dimensions of the different magnetic metals. He finds that iron, which first expands with the magnetizing force, soon reaches a maximum point, whence it retracts until it attains its original length; but, on still further increasing the magnetizing force, it contracts until it apparently reaches a minimum point, beyond which his means have not enabled him to proceed. His apparatus was so sensitive that he could read a variation of one hundred-thousandths of a millimetre. From *Nature* 37, 448; 8 March 1888.

Transposition of *copia* elements in *Drosophila*

SIR—Biémont *et al.*¹ present interesting new evidence for the movement of *copia* elements in the *Drosophila* genome, but I do not agree that transposition is necessarily the mechanism that explains this and similar^{2,3} observations.

The term should be reserved for the homology-independent insertion of transposon sequences into their target (often accompanied by a small duplication of the target site). Biémont *et al.* observed the appearance of new copies and the loss of existing copies at similar rates. The proposed reverse-transcription mechanism of transposition of *copia* and other retrotransposons does not require loss of the donor copy^{4,5}; rather, it is likely that loss of copies occurs most frequently by a homology-dependent process unrelated to transposition, namely LTR–LTR recombination (see figure). LTR–LTR recombination is very frequent in yeast^{6,7}, and has also been observed with *copia*⁸. The reverse reaction (conversion of a 'solo' LTR to an intact retrotransposon), also occurs at high frequency⁷.

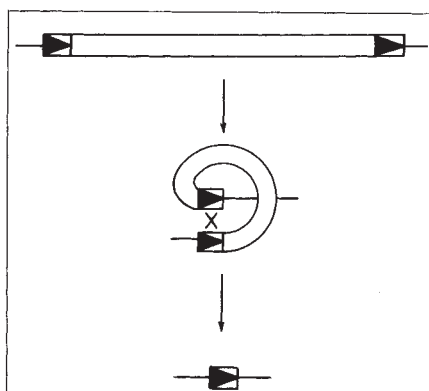
In situ hybridization techniques such as those used by Biémont *et al.*¹ and in related studies^{2,3} cannot reliably distinguish between the transposition and the homology-dependent movement of retrotransposons. It may well be that the events observed in these studies in fact reflect an increase in the activity of general homologous recombination and not of reverse transcription.

JEF D. BOEKE

Department of Molecular Biology
and Genetics,
Johns Hopkins University School
of Medicine,
725 N. Wolfe St.,
Baltimore, Maryland 21205, USA

1. Biémont, C., Aouar, A. & Arnould, C. *Nature* **329**, 742–744 (1987).
2. Gerasimova, T.I., Mizrokh, L.J. & Georgiev, G. *Nature* **309**, 714–716 (1984).
3. Gerasimova, T.I. *et al.* *EMBO J.* **4**, 3773–3779 (1985).
4. Mount, S. & Rubin, G.M. *Molec. cell. Biol.* **5**, 1630–1638 (1985).
5. Boeke, J.D. *et al.* *Cell* **40**, 491–500 (1985).
6. Winston, F. *et al.* *Genetics* **107**, 179–197 (1984).
7. Ciriacy, M. & Breilmann, D. *Curr. Genet.* **6**, 55–61 (1982).
8. Zachar, Z. *et al.* *Genetics* **111**, 494–515 (1985).

SIR—A recent paper by Biémont *et al.*¹ claims to demonstrate that transposition of the *copia* element of *Drosophila melanogaster* takes place in bursts. The evidence for this comes from *in situ* hybridization analysis of the polytene chromosomes of a strain (line 8) that has been maintained by brother–sister mating for 69 generations. There were a number of changes between generations 52 and 69 in the numbers and locations of *copia* elements in this line, with a gain of 11 and a loss of 8. Unfortunately, the data presented in the paper do not give a clear



LTR–LTR recombination. Boxed triangles represent LTR sequences; open box, the internal segment of a retrotransposon; and line, the flanking cellular sequences. One mechanism for the conversion of an intact retrotransposon to a 'solo' LTR reciprocal recombination at the LTR sequences, is indicated. Gene conversion may also lead to LTR–LTR recombination. The reaction, regardless of mechanism, is freely reversible.

picture of the genetic status of line 8 in generation 52. It seems that only one larva from this strain was analysed in this generation, and so the possibility cannot be excluded that it was still segregating for polymorphic *copia* insertions. The apparent change between generations 52 and 69 could thus simply be due to random changes in element frequencies at the chromosomal sites in question. The authors' earlier report on the distribution of the *mdg-1* element indicates that these lines were by no means homozygous after extensive inbreeding². Without a fuller characterization of their state in generation 52, no firm conclusion can be drawn concerning the mode of transposition of *copia*.

BRIAN CHARLESWORTH

Department of Biology,
The University of Chicago,
1103 E. 57th St.,
Chicago, Illinois 60637, USA

CHARLES H. LANGLEY

Laboratory of Genetics, NIEHS,
PO Box 12233, Research Triangle Park,
North Carolina 27709, USA

1. Biémont, C., Aouar, A. & Arnould, C. *Nature* **329**, 742–744 (1987).
2. Biémont, C. & Aouar, A. *Heredity* **58**, 39–47 (1987).

BIÉMONT *ET AL.* REPLY—The proposal of Boeke that LTR–LTR recombinations instead of transpositions are involved in our observation of a genome reshuffling for *copia* in one of our inbred lines¹, is based mainly on the fact that we observed simultaneously appearance of new copies and disappearance of existing ones. Because replicative transposition does not require loss of the donor element, Boeke concludes that transposition is not the main event in reshaping the genome.

Although we do not want to say that homologous recombinations do not exist in our lines, two facts argue against such an hypothesis in genome reshuffling. First, as the *copia* copy number is regulated in inbred lines, as recently reported (there is a compensation in copy number between the chromosome arms^{2,3}), then new insertions may be apparently associated with loss of existing copies so that the total copy number is maintained. Indeed, the line submitted to the genome reshuffling after more than 50 generations of sib crosses had 14 *mdg-1*, 14 *I*, 15 *copia* (it had 18 *copia* insertions before the reshuffling), and 14 *P* element copies¹; this line had the smallest copy number of mobile elements of all our 17 lines^{2,3}. Second, an "increase in the activity of general recombination" by LTR–LTR homology, as proposed by Boeke, should have led to movement of other mobile elements having two homologous LTRs at their extremities. No such simultaneous movements were observed in our line for *mdg-1*, a *copia*-like element¹. Boeke is right in saying that the *in situ* hybridization technique cannot distinguish between true transposition events and LTR–LTR recombinations, but this is not what we usually ask the technique to tell us. We used it to detect eventual movements of mobile elements in highly inbred genomes, and to localize the insertion sites precisely^{1–3}.

From their comments, Charlesworth and Langley seem to have understood that the conclusion of our experiment is that the usual mode of transposition of *copia* is by bursts. This is far from our own idea. What we really observed was a drastic change in *copia* insertion locations in one line (No. 8) between generations 52 and 69. Such drastic changes were not observed in other lines in which only a few new insertions or losses of sites were reported¹. The change in *copia* insertion pattern of line 8 was not the result of segregation for polymorphic insertions. Indeed, this line 8 was characterized by a low number of insertions for many elements³ and no polymorphism at all was observed for *mdg-1* during more than 80 brother–sister generations. At generation 52 there was no detectable insertion polymorphism for *copia* either (3 larvae were checked). Moreover, as reported in our paper¹, only a residual *copia* insertion polymorphism was observed at generation 69, suggesting that the pattern of *copia* insertions at this generation was not the result of continuing movement of the element every generation, contrary to the hypothesis of Charlesworth and Langley. That there was no additional change in *copia* locations in the 70th and 71st generations, and, as recently checked, in generation 80, strongly reinforces this conclusion. We are thus forced to involve a high rate of *copia* movements between generations 52

and 69. Note that whatever the mechanisms involved in this reshuffling, its causes remain to be determined.

C. BIÉMONT

A. AOUAR

C. ARNAULT

*Laboratoire de Biologie des Populations,
Université Claude-Bernard, Lyon 1
69622 Villeurbanne, France*

1. Biémont, C., Aouar, A. & Arnault, C. *Nature* **329**, 742–744 (1987).

2. Biémont, C. & Aouar, A. *Heredity* **58**, 39–47 (1987).

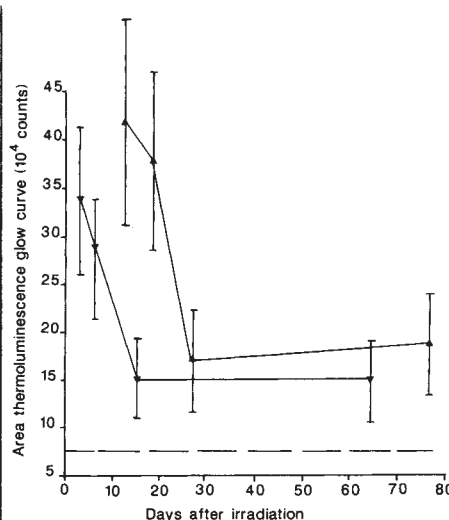
3. Biémont, C. & Gautier, C. *Heredity* (in the press).

Thermoluminescence in irradiated foodstuffs

SIR—The UK government has recently confirmed that a ban on the irradiation of foodstuffs will remain in force until suitable regulatory controls, including a simple method to detect whether food has been irradiated, can be devised. In 1983, the United Nations Codex Alimentarius Commission¹ of the Food and Agriculture Organisation/World Health Organisation stated that “despite the many investigations designed to detect physical, chemical and biological changes in foods subject to ionising energy, no satisfactory method for identifying food as having been irradiated has so far been developed”. Methods previously used in such investigations were found to be unsatisfactory for a variety of reasons², but thermoluminescence³ appeared to us to be worthy of investigation. We have measured the thermoluminescent response of seven different spices and found that most irradiated specimens could be clearly distinguished from unirradiated ones up to two to three months after irradiation.

The seven spices chosen for investigation were nutmeg, paprika, black pepper, cinnamon, cumin, turmeric and cayenne pepper. These are likely candidates for commercial irradiation and would receive doses of up to 10 kGy. Two samples of each spice received radiation doses of 5 kGy and 10 kGy, respectively, in a gamma-irradiation facility (cobalt-60 source) while a third sample served as an unirradiated control. The light output of samples from each group was subsequently measured with a photomultiplier tube while they were being heated from ambient temperature to 180°C at a linear heating rate of 2.5°C s⁻¹. A steady flow of nitrogen was used during heating to reduce smouldering.

We obtained simple, reproducible glow-curves; for example, the curve for turmeric is a single broad peak with a maximum at about 130°C, decreasing in height (fading) as the interval between irradiation and readout increases. Initial fading is rapid for all the spices (except nutmeg), but the thermoluminescent response persists well above that of



Integrated light output at different times after irradiation from cumin that had received radiation doses of 5 kGy (—▼—) or 10 kGy (—▲—), and from unirradiated cumin (---). Error bars are 1 s.d. (25 %) of values obtained by repetitive measurements.

unirradiated spices for over two months after irradiation. The results for cumin are shown in the figure. We obtained similar results for paprika, cayenne pepper, cinnamon, black pepper and turmeric, the highest response being from turmeric (twice that of cumin).

Nutmeg has the lowest response of all the spices investigated, with light output approximately 30 times less than that of cumin. Its response fades to that of unirradiated nutmeg after two and four weeks for the 5 kGy and 10 kGy-irradiated samples, respectively. Spices irradiated with 10 kGy initially show twice the thermoluminescent response of those irradiated with 5 kGy doses; however, further work is necessary to confirm whether the response is linear for a range of absorbed doses.

The errors obtained in this preliminary study are quite large but could be reduced by improving the heating and light measurement systems. It seems possible that this technique could be applied to a broad range of other foodstuffs and it therefore appears to be a promising candidate in the search for a simple test of whether food has been irradiated, as demanded by government, the food industry and public opinion.

T. F. MORIARTY

JENNIFER M. ODUKO

N. M. SPYROU

*Department of Physics,
University of Surrey,
Guildford,
Surrey GU2 5XH, UK*

1. Codex Alimentarius Commission, Vol. XV *General standard for irradiated foods* (FAO/WHO, Rome, 1983).

2. Deschreider, A.R. *Results of a European Research Programme: Proc 2nd Int. Conf. on the identification of irradiated foodstuffs* 305–325 (EEC, Luxembourg 1974).

3. Heide, L. & Bögl, W. *Fresenius Z. Analyt. Chem.* **320**, 682–683 (1985).

Techniques of pyramid-building in Egypt

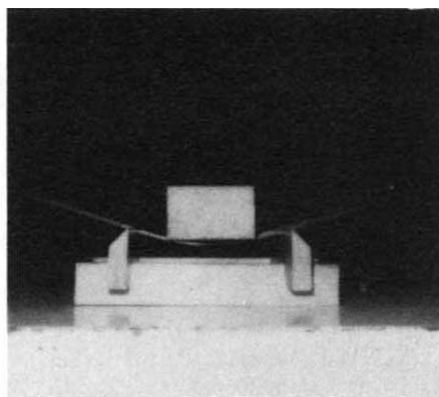
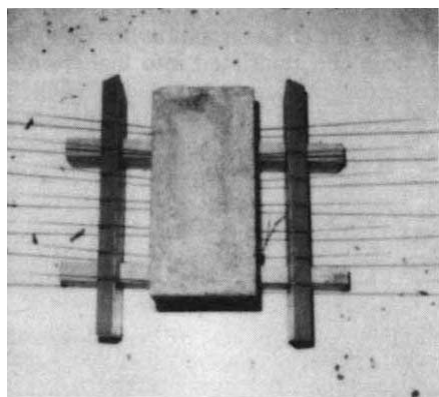
SIR—A principle illustrated in 5,000-year-old Egyptian art work has yielded a novel method of creating mechanical advantage. It has long been known that ancient peoples had the ability to lift and transport huge masses of stone. The Incan fortress walls, for example, contain blocks which weigh over 300 tons¹ and the monuments on the Giza plateau, which include the Great Pyramid, involved the transportation of over 11 million cubic yards of stone. In both cultures this was mysteriously accomplished without the wheel, the pulley or domesticated beasts of burden².

Considerable debate has been engendered by this mystery^{3–6}, and many knowledgeable authorities have concluded that ancient peoples may have had a method of lifting and transporting heavy weights that is unknown today.

In ancient Egyptian art, loads are commonly depicted being carried on poles⁷ and actual poles exist in museum collections. The evidence appears to indicate that ancient peoples routinely lifted and carried heavy weights by placing them on slender, bending poles, poles arranged symmetrically and often placed in parallel under a load. (Interestingly, this method of carrying things is still used in parts of the world today.)

I have built models consistent with these observations. A load represented by a cement brick 3½ inches wide, 7½ long and 2½ inches thick was placed on seven stainless steel rods 0.045 inches in diameter. The rods were evenly spaced under the brick and arranged perpendicular to the brick's long axis. These rods were supported on a wood frame consisting of two parallel wooden supports spaced 6½ inches apart. The rods were long enough to extend several inches to either side of the bearing supports (see figure). This construction establishes an equilibrium state between the bending rods and the load. Each rod supports only the share of the brick's weight which is acting directly above it. Lifting one rod end up and blocking it at a new height (¼ inch above the initial height, for example) causes the entire load to move up; one has added energy to the system, the rods that have not been moved straighten slightly and the rod elevated is bent more. Equilibrium must be maintained over the entire system. It can be observed that the centre of mass, with one rod end elevated, has moved up much less than ¼ of an inch.

If all the rod ends are lifted the same distance, one at a time, and blocked or shimmed up in an elevated position above the initial bearing point, the brick gradually rises. When the last rod end is elevated, the brick will have been lifted the same distance the individual rod ends



A model of the lifting device at Skidmore College. Using twelve oak poles, each $1\frac{3}{4}$ inches square and 14 feet long, a load of 2,600 pounds has been lifted. Left, top view; right, end view.

were raised. It is apparent that the distance the load is lifted is proportional to the number of rod ends moved. When all the rods have been lifted and blocked the system has returned to its original configuration. The brick, however, now exists in a position in space higher than it was initially. The blocking process may then be continued, lifting the load still further.

Consequently, for any one rod blocked up a distance s , the load is raised a distance close to $s \times 1/N$ where N is the total number of rod ends in the system. If n is the number of rod ends simultaneously lifted then the ideal mechanical advantage for symmetrical pole configurations will be N/n . If the rod ends are raised one at a time, the ideal mechanical advantage will be exactly N (14 for the above model). It can be observed that by placing extra rods under the brick the mechanical advantage of this pole system can be a matter of choice. The force of any pole end could easily have been designed to fall within the range that human muscles are capable of exerting. For example, if 100 pounds is applied to each pole end of some larger pole system possessing the mechanical advantage of the model, a load a little less than 1,400 pounds could be lifted.

This system can easily be confused with a support system composed of multiple second-class levers. Seven $\frac{1}{4}$ -inch wooden dowels may be substituted for the bending rods to observe the behaviour of such tools in this situation. Lifting up on the ends of the individual dowels, then changing the number of dowels supporting the brick and repeating the process, effectively demonstrates that no force-multiplying effect is to be gained by adding extra rigid supports. An attempt to lift any lever end results in force being applied to a large part of the entire load; unlike poles, levers do not bend to accommodate as individual supports are raised. It is revealing that levers such as these were never portrayed as being used to move and transport things in Egyptian art yet, paradoxically, they are among the first tools proposed by contemporary observers as pyramid-building implements.

Bamboo would be an ideal natural material for pole systems, although any support capable of bending will function as the model suggests. An ideal twentieth-century lifting system would probably be composed of carbon-fibre vaulting poles.

This evidence implies that it may have been possible that ancient peoples were aware of a force-multiplying mechanism that is not utilized in contemporary times. The device combined the weight-distributing principle of the suspension bridge with the energy-storing characteristics of the bow. The ancient Egyptians, for example, may not have needed to use ramps in order to lift the massive building materials used in their great constructions on the Giza plateau. More importantly, there may exist new contemporary applications of the principle as well⁸.

JOHN CUNNINGHAM

Skidmore College,
Saratoga Springs, New York 12866, USA

1. Protzen, J.-P. *Scient. Am.* 94–105 (February 1986).
2. Hamblin, D.J. *Smithsonian* 78–93 (April 1986).
3. Tellefsen, O. *Natural History*, 11 (November 1979).
4. Weeks, K.R. *Natural History*, 6 (November 1979).
5. I.E.S. Edwards, *Keeper Natural History*, 10 (December 1979).
6. Tellefsen, O. *Natural History*, 10 (December 1979).
7. Wilkinson, C.K. & Hill, M. *Egyptian Wall Paintings, The Metropolitan Museum of Art's Collection of Facsimiles* (Metropolitan Museum, New York, 1983).
8. US Patent Office patent applied for.

Avoiding jet lag again

SIR—The recent comments of Winfree¹, together with the anecdotal data from Woodruff² and Ford³, raise the issues of whether adjustment to a new time-zone can be promoted by behavioural changes (including meals and exercise) and whether this process of adjustment can begin during the flight itself. Woodruff also suggests that there is some value in correlating the recovery from the symptoms of jet lag with more objective measures of the adjustment of the body clock.

A chief problem is that jet lag is a 'rag-bag' of symptoms⁴ that seem to affect individuals in different ways and to different extents, although a detailed and systematic study remains to be carried

out. Inappropriately timed fatigue is one measure of jet lag that is used (if only because it is easy to assess) but other symptoms — difficulty in concentrating and gastrointestinal disorders, for example — might be of more importance to the traveller. Objective estimates of the process of adjustment of circadian rhythms to a time-zone transition have been made using a wide variety of variables; these range from tests of mental performance through measurements of deep body temperature, plasma cortisol and rapid-eye-movement sleep^{4–6}. Such objective measures have shown that the duration of the flight, and whether one is returning to, or leaving, one's home time-zone, are of less importance than the number of time-zones crossed and the direction of flight⁶. A detailed comparison between the rates of adjustment of such objective variables and the loss of the various symptoms of jet lag after a change in time-zone would make an interesting study.

The observations of Woodruff and Ford can be interpreted to indicate that adjustment to a new time-zone can be promoted (or at least the symptoms of jet lag can be minimized) by adopting the timing of life in the new time-zone as soon as possible. Existing evidence on this point confirms this idea⁶ but very little work seems to have been carried out. We have interpreted this kind of result⁷ as evidence that the cues that adjust the body's timing mechanism include the timing of our business, sleep, meals and leisure.

Attempting to start adjustment during the flight seems a natural outcome of the above considerations⁷. In practice, difficulties can arise in, say, a flight to the United States that starts from England in the early morning; travellers would be required to attempt to go back to sleep on the plane, having already had a full night's sleep. Similarly, a traveller journeying to the east has to contend with attempting to begin a night's sleep before he feels tired or with missing one night's sleep and staying awake for an unpleasantly long period of time. A short nap on arrival to his destination can reduce sleepiness, but can lead to lack of tiredness when the next night comes around. Whether rest in the dark would be an adequate substitute for sleep — and whether isometric exercises would be equivalent to the hamster's running wheel⁸ if local time at the traveller's destination indicated it was daytime — are further areas that still remain to be investigated.

As Winfree¹ says, the daily air-traffic provides a huge amount of potential material for use in investigating such problems. With the increasing awareness of the problem of jet lag on the part of the general public, perhaps some airline will realize that 'passenger care' might include attempts to minimize this concomitant

of air travel, and will become involved in studies such as those that have been mentioned.

D.S. MINORS
J.M. WATERHOUSE

*Chronobiology Group,
Department of Physiological Sciences,
University of Manchester,
Manchester M13 9PT, UK*

1. Winfree, A.T. *Nature* **330**, 311–312 (1987).
2. Woodruff, M. *Nature* **331**, 217 (1988).
3. Ford, B.J. *Nature* **331**, 309 (1988).
4. Holley, D.C. et al. *Effects of Circadian Rhythms Phase Alteration on Physiological and Psychological Variables: Implications to Pilot Performance* NASA Tech. Mem. No. 82177 (1981).
5. Aschoff, J. et al. *Chronobiologia* **2**, 23–78 (1975).
6. Klein, K.E. & Wegmann, H.M. *Significance of Circadian Rhythms in Aerospace Operations*. AGARDograph No.247 (NATO, 1980).
7. Minors, D.S. & Waterhouse, J.M. *Aviat. Med. Quart.* **1**, 9–26 (1987).
8. Mrosovsky, N. & Salmon, P.A. *Nature* **330**, 372–373 (1987).

Roaring and oestrus

SIR—Karen McComb's account of her interesting experiment on the effect of roaring on oestrus in hinds (*Nature* **330**, 648; 1987) is marred by the analysis of her results. To remove 'skewness' in the distribution of calving dates by log transformation (a liberty, at best); to 'adjust' them to take account of two other significant variance components and then plot as a cumulative percentage is over-egging the statistical pudding. In reports such as this where the amount of data is quite small, could it not be given raw, so those with a particular interest can transform it as they wish? After deducing from the plot some idea of what the actual calving dates might have been, I could suggest that the case McComb makes out is not as convincing as is implied by the statistical contortions she has used.

J. C. BIGNALL

*The Health Centre,
Newport, Dyfed SA42 0TJ, UK*

McCOMB REPLIES—Bignall's criticisms of my recent paper are mistaken. In that paper I plot cumulative percentage calving in three treatment groups to show that red-deer hinds exposed either to playback of recorded roaring, or to a vasectomized stag, calve earlier than control hinds (see my Fig. 1). I found these effects to be statistically significant in an analysis of variance. Bignall seems to believe that the plot of cumulative percentage calving (my Fig. 1) shows log-transformed values. This is clearly not the case, for a plot of the log-transformed (normalized) data would produce a symmetrical logistic (sigmoidal) curve. If Bignall had, as he suggests, "deduced from the plot some idea of what the actual calving dates might have been" by anti-logging the values, he would have obtained some very odd results indeed. The last calves would then have been born $> 2.7 \times 10^6$ years after the first.

Bignall implies that log transformation

is a misleading contortion of the data. But real data do commonly exhibit some skewedness, and under such circumstances it is usual to normalize data by log transformation before analysis of variance is performed to ensure that the assumptions of the analysis of variance model are not violated (Snedecor, G. W. & Cochran, W. G. *Statistical Methods*; Iowa State Univ. Press, Ames, 1967). Failure to use normalized data is, under such circumstances, far more likely to give misleading results.

Bignall also complains that I have statistically controlled for sire and sex effects, implying that the treatment effects shown in Fig. 1 might disappear if I did not do so. Analysis of variance reveals that both sire stag and a sire stag $\times$ sex-of-calf interaction have significant effects on the variance in calving date. As these confounding effects fell disproportionately across treatment groups (some hinds not completing the experiment) it was appropriate to control statistically for them when examining the effect of treatment. However, as I describe in my paper, the effect of treatment was significant before the confounding effects were removed.

KAREN McCOMB

*Large Animal Research Group,
Department of Zoology,
Cambridge CB3 0DT, UK*

Novel databases for molecular biology

SIR—Pabo¹ has highlighted the need for second-generation databases for molecular biology. One important task would be to organize non-structural data into accessible databases, which in itself would justify a major collaborative effort². Furthermore, structural information should be accessible in the same conceptual framework as non-structural data. Development of a uniform and coherent conceptual framework could be based on a systematic study of terms, concepts and cognitive structures we use to perceive, to interpret and to memorize macromolecular data.

The approach suggested here is based on the analysis of scientific communications and statements pertinent to structural data. A preliminary — and by no means impartial — survey of current papers on protein and DNA structure suggests that the conceptual machinery we use to describe macromolecular data is in fact simple and uniform.

For example, macromolecular structures, which are too complex to be memorized, are recoded as a simplified set of substructures related to each other by a few relationships. The constituent elements are in turn characterized in terms of size, composition and similarity to other known substructures. The relationships of elements are usually described as sym-

metries and vectorial distances. Primary protein and DNA structures, for example, seem to be translated into higher-order sequences of complex elements (such as cleavage sites or ligand-binding sites). These segments can be described in different ways, as there are several symbolic and parametric methods to represent sequences.

Structures of DNA and protein can best be visualized as hierarchical and colinear data-structures (represented by sequences of symbols, numbers or sub-files), on which the same types of operations (such as insertion, deletion, exchange, symmetry operations) can be defined. Primary sequence similarity is by far the most frequent tool for detecting and characterizing structural similarities, even though it is becoming apparent that distant homologies are more easily understood in terms of parametric representations³.

Even these simple concepts can be highly efficient if used in a generalized sense. For example, notions like a repetitive pattern of beta strands or amphiphilic helices can both be described as translational symmetries (in secondary structure and in hydrophobicity, respectively). Similarly, the statement 'a hydrophobic helix flanked by ionic residues' can be considered as a specific sequence of three elements which are defined in terms of composition.

A generalized framework of concepts and relationships abstracted from human thinking could enormously increase the performance of molecular-biology software. New categories identified by computer-based methods could be introduced to complement the known substructures. For example, linguistic methods have been used for the automatic definition of characteristic subwords and syntactic rules in DNA⁴. Moreover, artificial intelligence methods could be built up using these extended concepts, for automatic analysis of the available structural data and for building up new structural databases. Novel database structures could thus be the final result, rather than the beginning of this process.

Interpretation of macromolecular data seems to be ripe for artificial intelligence methods. Highly accurate data are available in abundance; concepts and terms used to describe these data are efficient and can be unequivocally defined in scientific terms; and our obvious inability to obtain an adequate overview of the ever growing body of information leaves us no other alternative.

SÁNDOR PONGOR

*Agricultural Biotechnology Centre,
2101 Gödöllő, PO Box 170, Hungary*

1. Pabo, C.O. *Nature* **327**, 467 (1987).
2. Davison, D. *Nature* **329**, 194 (1987).
3. Pongor, S. *Meth. Enzymol.* **154**, 450–473 (1987).
4. Brendall, V. J. *biomol. Struct.* **4**, 11–22 (1985).
5. Bocksteele, F. *Biochimie* **67**, 509–516 (1985).

Out of the ivory tower

John Pfeiffer

Beyond the Laboratory: Scientists as Political Activists in 1930s America. By Peter J. Kuznick. University of Chicago Press: 1987. Pp.363. \$29.95, £23.95.

ONCE upon a time scientists dedicated themselves to what one of them called "the unfettered pursuit of pure knowledge", living on relatively tiny budgets out of the public limelight (except of course for Einstein, who represented genius in the clouds). Fred Allen, a popular comedian in the early days of radio, poked fun at the profession in one of his weekly broadcasts. During the course of an interview with an actor impersonating a physicist, he asked why anyone would spend his life trying to smash atoms. The reply, delivered in a vaudevillean German accent, brought down the house: "Vell, some-day someone might vant half an atom".

Science's Ivory Tower days came to an end half a century ago. It took social upheavals of seismic proportions to jolt them at long last into the thick of twentieth-century politics, notably the Great Depression, Nazism on the rise and a world war in the offing. The extent of their naïveté is recounted in *Beyond the Laboratory* by Peter Kuznick, a historian at the American University in Washington, DC, who focuses on what happened during the 1930s in the United States. Ernest Lawrence acted firmly to keep his colleagues out of the mainstream, and "effectively banned" all discussions of social and political issues in the University of California Radiation Laboratory. Robert Oppenheimer first learned about the 1929 stockmarket crash "long after the event . . . Lawrence told me about it when we were on a walk".

Other scientists learned about the crash less casually. According to a 1933 estimate, some 2,000 chemists joined the ranks of the unemployed in the New York area alone. Many of those not fired put up with across-the-board pay cuts. The University of Washington, for example, established a sliding-scale policy: full professor salaries reduced 29 per cent, associate professor salaries 23 per cent, instructor salaries 15 per cent. What hurt most of all was the massive curtailment of federal funds for research. In their own way politicians turned out to be as naïve as scientists. The greatest politician of them all, Franklin Roosevelt, knew little

about science and cared less, at least in the beginning. His New Deal came down hard on researchers who saw their already reduced budget for fiscal year 1932–1933 slashed 60 per cent, from \$66 million to \$26 million.

A subsequent executive order seemed

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Franklin D. Roosevelt — no new deal for science.

to herald better times. Responding to arguments, mainly from members of his own cabinet, Roosevelt appointed an élite Science Advisory Board including Frank Jewett, John Merriam, Karl Compton and W.W. Campbell — presidents, respectively, of the Bell Telephone Laboratories, the Carnegie Institution of Washington, the Massachusetts Institute of Technology and the National Academy of Sciences. According to Kuznick, the move turned out to be "an exercise in futility", and that is putting it mildly. Nothing happened. The Board failed utterly to agree about new guidelines for federal support. More generous budgets would obviously mean more research projects and reduced unemployment. But, and a very big 'but' it was in those days, the danger of political interference with basic research and the free flow of information represented something to be avoided at all costs. Besides, some Board members felt that in the long run the Depression might not be all that bad, because the investigators hit hardest probably ranked among the least competent. It was hardly science's shining hour.

In any case, dissension within the ranks became increasingly irrelevant as the sheer pressure of events accelerated the decision process. An early sign of worse things to come was the rising tide of refugee scientists from Nazi Germany. Emergency committees found jobs for hundreds of them, over the objections of some Americans who believed that the native unemployed merited first consideration, a position often sullied by "overt feelings of anti-Semitism". The American Association for the Advancement of Science organized a series of "Science and Society" conferences "to investigate and present in a systematic and comprehensive way the effects of science and its applications upon society and upon human beings as individuals".

BBC Hulton Picture Library

The American Association of Scientific Workers, formed in 1938 along the lines of the British parent body and rather more activist than the AAAS, dedicated itself to raising the social consciousness of scientists. It worked with the press and publishers to expose the fallacies of racism and pseudoscience, and promote public understanding of science in general. A special AASW committee investigated the possible effectiveness of boycotting \$8 million worth of German optical equipment, technical publications and other items, speculating that the move might force large cutbacks in the manufacture of

Nazi armaments and perhaps help contribute to the eventual collapse of the Hitler regime.

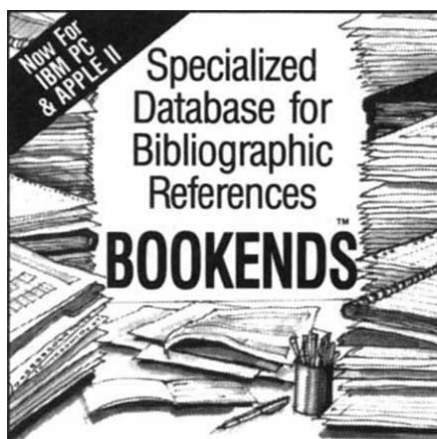
Not all AAAS and AASW members were as wholeheartedly anti-Nazi. A number of eugenically minded biologists looked with some favour upon Germany's initial effort towards 'racial betterment', a 1934 law establishing the compulsory sterilization of hereditary defectives (a category broadened a year later to include Jews). Far more prevalent at home and abroad, among scientists and non-scientists alike, was the bulwark-against-communism notion. Much of what was going on in Germany could be tolerated if not condoned, if only the Nazis and Soviets, like rival gangster mobs, would save the forces of law and order a great deal of trouble by killing each other off. That hope, and a good many others, died in the summer of 1939 with the signing of the Nazi-Soviet pact, which did more damage to ideals and ideologies in a briefer period than any previous act of diplomacy.

These days, scientists are more sophisticated than they were in the dim and distant 1930s. They are certainly more

active in public affairs, more closely associated with government and industry, more publicity-conscious. But the fear of communism and government interference is still with us. Scientists are more deeply divided than ever, with dedicated spokesmen on both sides of crucial issues — for and against nuclear disarmament, research on Star Wars, the international exchange of the results of basic research.

Kuznick's book documents a crucial period in the history of science. I wish it had been more carefully edited, however. A number of important individuals are inadequately identified or not identified at all. Also, the journal in which this review appears is mistakenly referred to as the "official organ" of the British Association for the Advancement of Science. *Nature*, needless to stress, is notably independent. Kuznick assumes an above-the-battle stance. He has virtually nothing to say about the present-day implications of his study, presumably on the grounds of letting the facts speak for themselves (a cop-out, because they never do). As far as the future is concerned, however, any serious analysis of science in our times will draw heavily on the record abundantly documented in *Beyond the Laboratory*. □

John Pfeiffer, PO Box 273, New Hope, Pennsylvania 18938, USA, is a science writer and journalist, and is currently at work on a book about the cultural evolution of our times.



Bookends allows you to quickly put your finger on any journal, any paper, any book, any specific information contained in your voluminous files.

Bibliographies in any journal format are quickly and easily produced. Editing can be performed with the ease of a word processor.

The *New England Journal of Medicine* states in its issue of January 12, 1984 that Bookends "addresses two widespread problems: first, the challenge of cataloging all the information we read; second...the need to convert bibliographic information into the format demanded by each...journal. Both problems are solved here in an excellent way."

Bookends Extended is available with above features plus 80-column display, more memory, and ability to handle unlimited references.

Bookends, the reference management software system is available for Apple computers for \$124.95. Bookends Extended—Apple or IBM \$149.95.



IBM is a trademark of International Business Machines Corp. Apple is a trademark of Apple Computer, Inc.

Sensible Software, Inc.

335 East Big Beaver, Suite 207,
Troy, MI 48063, (313) 528-1950

Epidemic story

Peter Newmark

And the Band Played On: Politics, People and the AIDS Epidemic. By Randy Shilts. *St Martin's Press, New York/Viking-Penguin, London: 1987-1988. Pp. 630. Hbk \$24.95, £15.95; pbk £8.95.*

PUBLISHED late last year with a good deal of success in the United States, *And the Band Played On* is now available in Britain, where AIDS — the subject of the book — will never become either the medical or the political problem that it has been in San Francisco, whence, in a sense, this book emerges.

Because of its very large homosexual community, San Francisco was one of the two US cities where AIDS first took hold. As the AIDS reporter of the *San Francisco Chronicle* since 1982, Randy Shilts, author of the book, is able exhaustively, and sometimes exhaustingly, to recount the agonizing dilemmas the disease posed for the city's homosexual community and politicians — groups that were by no means mutually exclusive.

For the homosexual community the dilemma centred on the bathhouses, where years of repression were vented in excessive and chillingly anonymous homosexual acts. Although it became increasingly evident that the bathhouses were a major cause of the spread of AIDS, the curtailment of these symbols of hard-won homosexual freedom was unthinkable to the less responsible segments of the community. For the politicians the dilemma was much more mundane: enforced closure of the bathhouses would have endangered their support from a community that had gained considerable political muscle.

While Shilts reserves some of his most bitter comments for those who refused to take action in San Francisco, his sword sweeps over a broader battleground. In what sometimes strikes a self-congratulatory tone, he condemns the rest of the nation's media for ignoring AIDS until it killed Rock Hudson and then began seriously to spread outside homosexual communities. There is clearly some force to this analysis. But it is surely naive of a newspaperman, of all people, to suggest that it was simply because homosexual matters were deemed to be irrelevant to readers that the slowly emerging AIDS epidemic at first received far less media coverage than the 1976 outbreak of Legionnaire's disease. The difference is that 29 deaths in short order are 'news', while sporadic deaths are not.

Another target for Shilts is the directors of blood banks. On the evidence he provides, there is little doubt that they were far too slow to take steps, before the

availability of antibody tests, to prevent blood being drawn from high-risk individuals — a mistake that is still costing lives. In addition he condemns federal agencies for not providing sufficient money at an early enough stage for AIDS research, and for covering up internal recognition of this shortfall — claims that are supported by material obtained under the Freedom of Information Act.

Less convincingly, Shilts attacks the medical establishment for being slow to investigate AIDS therapy. Take the case of HPA-23, the drug that took Rock Hudson to Paris, from where he returned close to death but in a blaze of publicity. It is hardly surprising that desperate AIDS patients would take the expensive trip to Paris in pursuit of a drug that, on anecdotal evidence, offered some hope. But it is equally unsurprising and, indeed, reassuring that the medical establishment would hold to its view that there is simply no substitute for the time-consuming procedure of a rigorous clinical trial to determine whether a drug is better than useless. After all, HPA-23, like a score of other drugs that have been rumoured to cure AIDS, saved no lives.

US scientists are the final target. It is a pity that on the first of rather few occasions on which Shilts dips his toe into scientific waters he makes a fool of himself: "... he had started culturing lymphocytes ... in a special cell line Gallo had developed called interleukin-II". But science is essentially peripheral to the main focus of attack. With a good deal of justification, Shilts claims that US scientists were too slow to recognize the discovery at the Pasteur Institute of the virus that causes AIDS, and then too concerned with their own stake in matters for the good of the swiftest possible progress of the research.

Events in the book are recounted chronologically. In unabashed journalistic style, the author interleaves his account of the politics of AIDS with progressive details of a clutch of individuals involved in one way or another in the unfolding story. Many are San Franciscan homosexuals, whose almost inevitable infection with the AIDS virus and progression to disease and, in several cases, death is a poignant reminder of the consequences of the combination of inertia, indecision and inexcusable delay that characterized the city's response to the epidemic.

The book is at its best when dealing with events in San Francisco, the eye of the storm. Distant events viewed through the storm clouds become distorted and are sometimes superficially treated. But the book deserves a wide reading outside the United States where it may still not be too late to prevent some of the same mistakes being made again.

Peter Newmark is Deputy Editor of *Nature*.

Journey to the centre of the Earth

Don L. Anderson

The Earth's Core, 2nd Edn. By J.A. Jacobs. *Academic*: 1987. Pp. 416. £40, \$65.

THE study of the Earth's core might seem to be as academic as the study of a black hole or a pulsar. If it were not for the core, however, compasses would not work, we wouldn't know about plate tectonics and we probably wouldn't even have evolved to worry about such things. The core is a sloshing ball of molten metal, half the radius of the Earth, in which is suspended a sphere of solid iron. It is the source of the magnetic field which shields us from harmful emissions from outer space. The liquid outer core and solid inner core reveal their secrets through their effects on seismic waves, the magnetic field and the motions of the Earth.

This second edition of Jacobs's book, first published in 1975, comes at a time of rapidly expanding interest in the core and information about it. It therefore serves as an update, or starting point, for a rapidly moving field rather than as the final word. The author manages in about 400 pages painlessly to cover much of the ground expected in an introductory survey course. The core turns out to be an exciting place.

The 'facts' concerning seismology,

geomagnetism, palaeomagnetism, Earth rotation and high-pressure laboratories are remarkably up to date and are sprinkled throughout the book as befits their diversity. The core cannot be studied in isolation and particular attention is paid to the core-mantle boundary and D'', the layer at the base of the mantle. Jacobs wisely uses Chapter 1 to describe the general physical properties of the Earth and writes an admirable introduction to the physics of the Earth's interior, including the core. He then describes the various theories and constraints on the origin and thermal regime of the core.

The reader who has never worried much about the core will be surprised that some 500 papers are mentioned in the first three chapters. Another 300 are drawn upon in the longest chapter, which covers the Earth's magnetic field. Other chapters treat the constitution of the core and the cores of other planets, and have their own extensive bibliographies. The large number of up-to-date references covering all aspects of the subject and related studies is an important feature of the book.

Jacobs has done an exceptional job of pulling together a huge amount of information and making it both interesting and understandable. Although the book is about the core, it is also a good place to learn about planetary interiors and how geophysics is done. □

Don L. Anderson is Professor of Geophysics and Director of the Seismological Laboratory at the California Institute of Technology, Pasadena, California 91125, USA.

Evolving editions

Joseph G. Gall

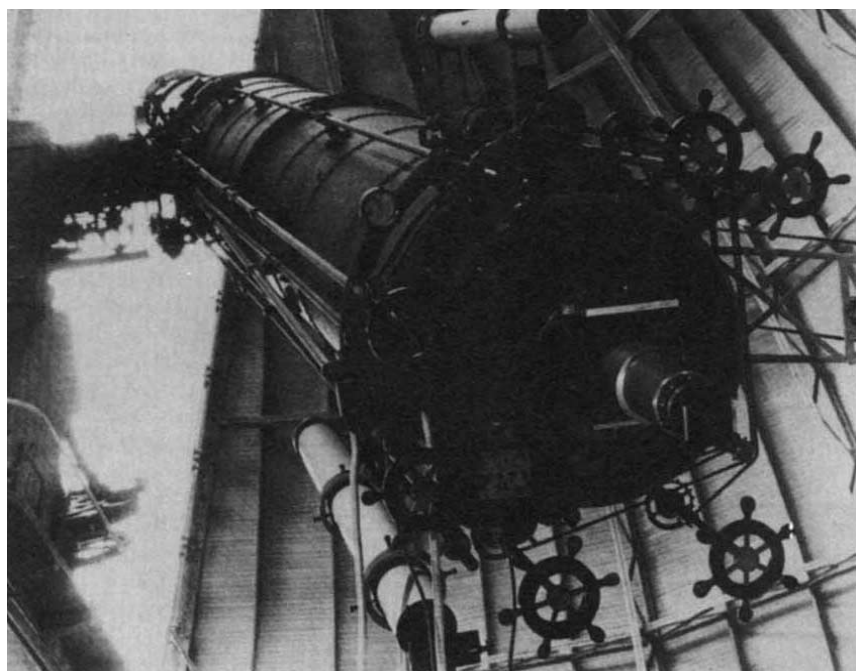
Molecular Biology of the Gene, 4th Edn. Vols I and II. By James D. Watson, Nancy H. Hopkins, Jeffrey W. Roberts, Joan Argetsinger Steitz and Alan M. Weiner. *Benjamin/Cummings*: 1987. Pp. 1,189. Vol. I \$39.95, £21.95. Vol. II \$29.95, £21.95.

JIM Watson's *Molecular Biology of the Gene* was first published in 1965. The book brought together for the first time the new prokaryotic genetics and the new biochemistry of nucleic acids and proteins, and it was enormously popular. The success of the first edition was a tribute to Watson's intellectual synthesis, helped greatly by his direct and lively presentation and the skilful use of *Scientific American* style diagrams and figures. The second and third editions of 1970 and 1976 required enlargement to accommodate much additional information, but they were completed before the revolution in recombinant DNA technology and DNA sequencing.

The fourth edition comes when every aspect of cell and molecular biology has been touched by these methods and as fresh information pours in at an ever-increasing rate. There are now five authors instead of one, the overall number of pages has more than doubled, and there are two volumes. The 28 chapters are wide-ranging, including not only the traditional subject matter of prokaryotic molecular biology, but extensive discussions about eukaryotic cells and multicellular organisms.

The first volume covers the general principles underlying gene structure and function in both prokaryotes and eukaryotes, and is intended to stand alone as the text for a beginner's course in molecular biology at the university level. Only the first few chapters, which deal with the elements of genetics and biochemistry, are recognizable from the previous edition; the rest are new or have been completely rewritten by Jeffrey Roberts and Joan Steitz. The second volume is much smaller, although in the preface we are promised that it will be expanded in subsequent editions. Written primarily by Nancy Hopkins and Alan Weiner, it concentrates on cellular differentiation and development, eukaryotic viruses, cancer and the origins of life.

This is an extraordinarily well-done project, clearly and concisely written, pleasing to the eye, and full of basic information on an incredible array of topics. A glance over the chapters shows how far we have come in understanding the molecular structure of genes and the processes they control — in every respect



Far-sighted — the largest refracting telescope in the world, at the Yerkes Observatory, University of Chicago. The telescope, finished in 1897, has an objective lens with a diameter of 102 cm. The picture is taken from *Fundamental Astronomy* edited by H. Karttunen et al., an introductory text for 1st and 2nd year university students as well as for serious amateurs. Publisher is Springer-Verlag, price is DM128. Nature's annual review of textbooks appears in next week's issue.

Molecular Biology of the Gene is a celebration of the crowning achievements of biology during the second half of the twentieth century.

So what is left to do? What problems remain for the fifth edition? Since the end of the nineteenth century biologists have sought a unified theory explaining what is special about the organization of living matter, how that organization is inherited, how it develops within the individual and how it evolves with time. These volumes display the great success of molecular biology in clarifying the first two areas. If the key to development is differential gene expression, as many believe, progress in the third area should now be rapid, if not explosive. Each day we hear exciting new information about gene regulation, how proteins interact with DNA to

turn genes on and off or up and down, and how growth factors, hormones and other extrinsic molecules regulate this activity.

In the fourth area — evolution — molecular biology will provide new arguments about prebiotic evolution and the origin of cells; and as DNA sequencing becomes ever easier, evolutionists will infer much about the past history of genomes from their present-day sequences. Thus we can confidently predict larger sections on developmental biology and genome evolution in the next edition of *Molecular Biology of the Gene*. In the meantime, these two volumes give us much to feast upon. □

Joseph G. Gall is American Cancer Society Professor of Developmental Genetics, Carnegie Institution, 115 West University Parkway, Baltimore, Maryland 21210, USA.

Instruments of change

Carole Stott

Early Scientific Instruments: Europe 1400–1800. By Anthony Turner. *Philip Wilson-Sotheby's/Harper & Row*: 1987. Pp. 320. £55, \$100.

The Divided Circle: A History of Instruments for Astronomy, Navigation and Surveying. By J.A. Bennett. *Phaidon-Christie's/Salem House*: 1987. Pp. 224. £45, \$75.

THE art world is becoming interested in scientific instruments. Auction prices haven't yet reached the impressionist level; the most coveted pieces are guaranteed only to reach a six-figure sum. But the dealers and auction houses have a glint in their eye and can see the way forward. Sale catalogues describe the material and mode of manufacture of instruments, rather than their function. The fact that a particular device would be pretty well useless and was only produced to adorn a gentleman's study does not lower the price. The more modern the instrument, the cheaper it is, not only because old-equals-value, but because the instruments of the nineteenth and twentieth centuries are functional and look functional.

Scientific instruments are not as old as science itself, and it was only in the sixteenth century that they began to be made in any number. These two volumes have both been published by leading auction houses which are keen to open up the subject to new customers, collectors who are equally capable of raising auction prices and affording expensive books.

Between them, the authors have 40 years of experience as historians, collectors and custodians of eminent museum collections. Their books are lavishly illustrated, drawing examples almost ex-



Divide and rule — Jesse Ramsden holding the old tools of his trade, but with his arm on the new dividing engine which drastically reduced the time and cost of graduating a scale.

clusively from European collections, and their choices of material reflect the geographical areas of original production and of present-day interest. Turner takes his examples from private and continental collections, Bennett is more firmly based in Britain. The result is that there is little overlap between the books though both provide excellent catalogues of extant instruments. The catalogues, however, are a by-product of the authors' aims, which are to describe the instruments and their development against the intellectual and social background of the time.

Turner's was perhaps the more daunting task. He has chosen to cover all types of scientific instrument, though it soon becomes clear that medical and associated instruments and chemical apparatus are excluded. But he does include many types which are not to be found in Bennett's book, the telescope, microscope, barometer and thermometer for example. Some of these instruments, however, are

introduced as if the reader knows already exactly what they are and how they work. Thus there is no clear explanation of the style or mode of construction of a Gregorian telescope, whereas the astrolabe is dissected, laid out in its parts and labelled thoroughly. This latter approach could have been used effectively throughout the book. Further evidence of Turner's assumption of scientific sophistication on the part of his audience is the sequence of four photographs of hair hygrometers, three of which lack their essential hairs and are therefore completely unusable.

Bennett concentrates on the angle-measurement instruments used in astronomy, navigation and surveying. He looks at three historical periods in the course of the book: the early modern period, and then the eighteenth and nineteenth centuries. The technique is the same, whatever the subject or period. The author lucidly describes the state of the science and the social context of the time. The leading advocates, the books of influence and the instruments themselves are discussed alongside photographs of extant examples. Nor are the instrument makers and their craft forgotten. Bennett tells the story of the pre-eminence of the English maker in the eighteenth century, men such as Dollond, Ramsden and Troughton, all Royal Society Copley Medal winners, who by the time of the Great Exhibition in 1851 had relinquished much of their position and trade to European manufacturers.

There it ends, however. We are told very little of how any of the instruments work and of the scientific principles behind them. The only attempt at this is in the first few pages of the book, where there is an explanation of right ascension and declination, but without any accompanying diagram and so conjuring up the image of a Frenchman trying to describe a spiral staircase with his hands in his pockets. Matters of instrumental accuracy and performance are also ignored. How easy was it to divide a circle? And how do a vernier scale and Ramsden's dividing engine, which "was to revolutionise the manufacture of portable instruments", actually work? We are left in the dark on such questions.

Both volumes contain an enormous amount of well-researched and well-presented information. Turner's notes and both bibliographies are excellent springboards for further investigation. Books on scientific instruments are scarce, and these two are valuable additions to the library of any serious collector and dealer. But there is plenty of room for authors to tackle the fascinating subject of the science behind the hardware. □

Carole Stott is Curator of Astronomy at the Old Royal Observatory, National Maritime Museum, Greenwich, London SE10 9NF, UK.

Dynamic models in behavioural and evolutionary ecology

Alasdair Houston*, Colin Clark^{†¶}, John McNamara^{‡¶} & Marc Mangel[§]

* Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

† Institute of Applied Mathematics, University of British Columbia, Vancouver V6T 1W5, Canada

‡ School of Mathematics, University of Bristol, Bristol BS8 1TW, UK

§ Departments of Agricultural Economics, Entomology and Mathematics, University of California, Davis, California 95616, USA

The maximization of fitness is often used to analyse the action of natural selection on the life history of animals but short periods of behaviour receive ad hoc treatment. This article describes a dynamic, stochastic model for analysing behaviour in terms of the maximization of fitness.

THE Darwinian concept of evolution by natural selection¹ has led to the introduction of optimization methods in biology²⁻⁴. The basic idea, as illustrated by life history theory⁵⁻⁷, is that behaviours have evolved in a way that maximizes some measure of reproduction (usually called 'fitness'). In general, life history theory deals with relatively gross decisions about allocation, such as yearly reproductive effort, and is not well suited for the description of fine-scale decisions, such as when a bird should remove an empty egg shell that might attract predators to the nest⁸. The classic work of ethologists such as Tinbergen and Lorenz showed that such fine-scale behavioural traits are also subject to the principles of natural selection. Here we describe mathematical methods that can be used to treat behaviour from an evolutionary standpoint by describing a framework for dynamic modelling of behavioural decisions in the context of the life history of an animal^{9,10}.

Earlier attempts to apply the principle of natural selection to fine-scale decisions typically maximized a simple surrogate currency instead of fitness¹¹. For example, most models that come under the heading of optimal foraging theory use the rate of energy gain as a currency, and assume that maximizing this rate is equivalent to maximizing fitness. Although these models sometimes can predict short term behaviours¹², they have limitations such as not being able to compare predation and starvation risks, not considering sequences of actions, and ignoring the state of the animal and the information it has about the environment. To include these aspects, a dynamic, state-variable approach to the analysis of behavioural decisions is required. For example, consider a fish in a lake which contains a number of different habitats¹³. If the potential food intake and risk of predation vary between habitats, how would it judge which was the better foraging habitat? Ideally, an optimization model should be able to explain both short-term (daily) movements between habitats and long-term habitat shifts over the lifetime of the fish. Thus we are concerned with both entire life histories and shorter periods in the life of an animal that involve detailed choices.

The four components of the framework we propose are a set of variables characterizing the state of the animal, a set of actions that the animal can perform, dynamics which specify the relationship between actions and subsequent states, and a state-dependent reward function that specifies future reproductive success in terms of the state of the animal at the end of a relatively short interval. To illustrate the framework, we begin with a simple paradigm for habitat selection⁹, then provide

examples concerning the behaviour of a number of different organisms.

A simple model

Imagine an animal that can choose one of n 'habitats' (H_i) in which to forage on a particular day t . Each habitat H_i is characterized by three parameters: β_i , probability of death due to predation, Y_i , energetic value of food items, λ_i , probability of discovering a food item.

The most interesting situation arises when the riskier habitats also happen to be the most productive. The question then is, assuming that evolution has led to the selection of 'optimal' foraging behaviour, which habitat would we expect the animal to choose?

To answer this question we introduce a state variable, X_t , representing the animal's bodily energy reserves at the start of day t . Thus energy reserves are increased by daily food consumption Z_t , and decreased by daily metabolic expenditure, denoted by α , so that reserves at the start of day $t+1$ are given by the equation:

$$X_{t+1} = X_t - \alpha + Z_t \quad (1)$$

Z_t is a random variable whose distribution is determined by the animal's choice of habitat. If habitat H_i is chosen, then Z_t equals Y_i with probability λ_i , and equals zero with probability $1 - \lambda_i$.

Energy reserves cannot be negative, and are bounded by the animal's capacity Cap:

$$0 \leq X_t \leq \text{Cap} \quad (2)$$

If energy reserves decline to a low level, the forager is likely to die of starvation; we simplify this situation by assuming that the animal dies if $X_t = 0$.

This framework now allows us to compare, in terms of fitness, the benefits of food consumption with the risks of predation. Death may occur either from starvation or predation. If, for the moment, we limit attention to behaviour during a nonbreeding phase in the animal's life, it may be reasonable to identify fitness with the probability of surviving until the end of the nonbreeding period. We then adopt the hypothesis that an optimal behavioural strategy is one that maximizes the forager's probability of surviving from $t=0$ to $t=T$, where T denotes a fixed horizon, corresponding to the onset of breeding activity. (This simple paradigm can be extended in many directions, one of which includes breeding activities.)

We assume that the animal 'knows' the parameter values β_i , Y_i , λ_i , characterizing each habitat H_i , and that it can assess its

¶ To whom correspondence should be addressed.

1. Derivation of the stochastic dynamic programming equation

Recall that $J(x, t, T)$ is the maximal probability of survival from day t to day T given that $X_t = x$. Hence, when $t = T$, $J(x, T, T) = 1$ as long as $x > 0$. For values of $t < T$, we proceed as follows. To survive from day t to T , the organism must first survive day t and then survive from day $t+1$ to T , taking into account changes in the state variable that occur on day t .

Suppose that the organism chooses habitat H_i on day t . If the animal avoids predation on this day, its energy reserves at the beginning of day $t+1$ can have one of two values, $x - \alpha$ and x'_i , where x'_i is the minimum of $x - \alpha + Y_i$ and Cap. These values correspond to finding no food and finding a food item of energetic value Y_i and occur with probabilities $1 - \lambda_i$ and λ_i , respectively. The corresponding maximal probabilities that the animal survives from day $t+1$ to T are $J(x - \alpha, t+1, T)$ and $J(x'_i, t+1, T)$ respectively. Thus, conditional on the animal avoiding predation on day t , its probability of surviving until T is

$$(1 - \lambda_i)J(x - \alpha, t+1, T) + \lambda_i J(x'_i, t+1, T).$$

Taking into account predation risk on day t the survival probability is

$$P_i(x, t, T) = (1 - \beta_i)[(1 - \lambda_i)J(x - \alpha, t+1, T) + \lambda_i J(x'_i, t+1, T)].$$

Because $J(x, t, T)$ is the maximum probability of surviving from day t to day T , we simply maximize P_i over habitats to find $J(x, t, T)$, which gives equation (4) of the text.

energetic state X_t on each day t . The computation of the optimal strategy for selection of habitat can be carried out using the method of stochastic dynamic programming. Details are given in Ross¹⁴, see also Box 1. Let $J(x, t, T)$ denote the probability of survival from day t until day T , given that $X_t = x$, and assuming that the animal uses the best strategy. We then have

$$J(x, T, T) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \end{cases} \quad (3)$$

(the animal is alive on day T if and only if $x > 0$).

It is shown in Box 1 that in general $J(x, t, T)$ satisfies the equation

$$J(x, t, T) = \text{maximum } P_i(x, t, T) \quad x > 0 \quad (4)$$

where

$$P_i(x, t, T) = (1 - \beta_i)[\lambda_i J(x'_i, t+1, T) + (1 - \lambda_i)J(x - \alpha, t+1, T)], \quad (5)$$

$$\text{and where} \quad x'_i = \min(x - \alpha + Y_i, \text{Cap}) \quad (6)$$

Equation (4) is known as the stochastic dynamic programming equation.

If the values of the model parameters are known, then J can be found by an iterative process called backward induction. We first take $t = T - 1$. Because of equation (3), all expressions on the right side of equation (5) are known. Thus $P_i(x, T - 1, T)$ is known for each choice of habitat H_i and $J(x, T - 1, T)$ can be found from equation (4). The value i^* of i that yields the maximum in equation (4) specifies the optimal habitat H_{i^*} . In general i^* depends on x . Having computed $J(x, T - 1, T)$ as above, one can then use the same process to compute $J(x, T - 2, T)$, and subsequently $J(x, T - 3, T), \dots, J(x, 0, T)$. The computations are easily coded for automatic computation, and can be performed very rapidly on a desktop microcomputer. The procedure is illustrated in Box 2.

The qualitative predictions are that (i) foragers will accept greater risks of predation in habitats with higher food availability, and (ii) a forager will accept greater predation risks when hungry than when it is well-fed. This provides an explanation for the decrease in feeding rate as total intake increases¹⁵. Such

2. Illustration of backward induction

We take the upper limit on reserves, Cap, to be 20 and consider three habitats, with $\beta_1 = 0$, $\beta_2 = 0.01$, $\beta_3 = 0.05$, $\lambda_1 = 0.4$, $\lambda_2 = 0.6$, $\lambda_3 = 0.8$ and $Y_1 = Y_2 = Y_3 = 2$, $\alpha = 1$. Thus as the predation risk in a habitat increases, so does the mean net gain. The table gives $P_i(x, t, T)$ for a range of x and t . (Top entry, P_1 ; middle entry, P_2 ; bottom entry, P_3). To compute these values, start at the right-hand column, where $J(x, T, T)$ is given by equation (3). At time $T - 1$, $P_i(x, T - 1, T)$ is given by equation (5), and $J(x, T - 1, T)$ is equal to the maximum value of P_i . This value is underlined. It can be seen that it is optimal to choose habitat 3 only if reserves are one. Calculating $J(x, T - 1, T)$ for all x shows that habitat 1 is chosen for all $x \geq 2$. The values of $J(x, T - 1, T)$ can be used to find $P_i(x, T - 2, T)$ and hence $J(x, T - 2, T)$ etc. Using this procedure it can be shown that, for $T - t$ sufficiently large, the optimal policy has the form 'choose habitat 3 if $1 \leq x \leq 2$, choose habitat 2 if $3 \leq x \leq 11$ and choose habitat 1 if $12 \leq x \leq 20$ '. This is the long-term policy for maximizing survival in this environment.

In addition to finding the optimal policy, it is possible to calculate the probability that an animal following it will be in any given state. From this, the probability that a given habitat is selected can be found. In this example, the probabilities of selecting habitats 1 to 3, respectively, are 0.526, 0.471 and 0.003. Thus in a large population of animals, the percentage of animals in these habitats would be 52.6, 47.1 and 0.3.

x		$J(x, t, T)$ is underlined				$J(x, T, T)$
		$t = T - 4$	$t = T - 3$	$t = T - 2$	$t = T - 1$	$t = T$
4	$P_1(x, t, T)$	0.9713	1.0000	1.0000	1.0000	1
	$P_2(x, t, T)$	0.9711	0.9900	0.9900	0.9900	
	$P_3(x, t, T)$	0.9409	0.9500	0.9500	0.9500	
3	$P_1(x, t, T)$	0.9426	0.9426	1.0000	1.0000	1
	$P_2(x, t, T)$	0.9521	0.9521	0.9900	0.9900	
	$P_3(x, t, T)$	0.9318	0.9318	0.9500	0.9500	
2	$P_1(x, t, T)$	0.7933	0.8560	0.8560	1.0000	
	$P_2(x, t, T)$	0.8378	0.8950	0.8950	0.9900	
	$P_3(x, t, T)$	0.8542	0.9440	0.9440	0.9500	
1	$P_1(x, t, T)$	0.3618	0.3618	0.4000	0.4000	1
	$P_2(x, t, T)$	0.5372	0.5372	0.5940	0.5940	
	$P_3(x, t, T)$	0.6873	0.6873	0.7600	0.7600	
0	$J(0, t, T)$	0	0	0	0	0

predictions are hardly surprising, but one cannot expect to derive deep insights from such a simple model. The intuitively appealing prediction that foraging behaviour should be affected by the forager's current state of hunger is one that does not arise from classical foraging models. Box 2 also shows that quantitative predictions can be made about the long-term distribution of animals among habitats.

The dynamic modelling approach offers several advantages over previous methods. First, the use of dynamic state variables means that models can be made more realistic and biologically meaningful. Constraints on variable values and rates are easily included—indeed they can hardly be omitted. The models can be made time-dependent and fully stochastic, realistically reflecting environmental conditions. Alternative behavioural choices (for example, foraging, resting or reproductive activities) can be treated simultaneously in a unified way^{9,10,16}. Evolutionary fitness can be modelled in a direct manner, by including reproduction as well as survival. In fact, this approach is a generalization of classical life history theory^{5-7,16}, but, by working with a representation of the animal's state, a better insight into the nature of life-history tradeoffs is obtained¹⁶. Moreover, the approach enables us to calculate the loss in expected future reproductive success that results from adopting a suboptimal action¹⁰, referred to as the canonical cost of an action¹⁰. These costs can be used to assess the robustness of the conclusions of

a model and the selection pressure for a given behavioural strategy^{10,17}.

In short, the stochastic programming approach to behaviour modelling is extremely flexible, and can be used to help organize and understand a wide variety of both field and experimental data; some recent applications are described below.

Foraging in lions

The African lion is the only social member of the cat family¹⁸. Lions live in prides typically consisting of up to 18 related adult females plus offspring and unrelated adult males. Female lions from a given pride hunt prey in cooperative groups ranging from one to eight, depending on prey size and other circumstances.

Caraco and Wolf¹⁹ analysed data collected in the Serengeti by Schaller²⁰ and calculated average daily food intake per lion as a function of hunting group size. In all cases, groups of size $n = 2$ maximized the average individual food intake. For large prey such as zebra and wildebeest, however, the average number of lions observed feeding at kills ranged from four to eight. Caraco and Wolf calculate that these large groups would experience up to 50% reduction in food intake, compared with groups of two. When hunting small prey such as Thomson's gazelle, on the other hand, lions either hunt individually or in small groups of two or three.

Packer¹⁸ has re-examined the results of Caraco and Wolf and concludes that "there are no good data showing that cooperative hunting is in fact beneficial to individual lions" in terms of average rate of food intake. Instead, Packer suggests that lion sociability is primarily an adaptation to the opportunity for scavenging from conspecifics in a region where both lions and their prey occur at high densities. When food items are large and infrequently obtained, variance in the amount of food acquired may have as great an influence on fitness as does the average feeding rate. By increasing the frequency of kills and decreasing the food per hunter per kill, group foraging generally decreases the variance in food intake.

To assess the implications of group foraging in greater detail, Clark²¹ developed a dynamic model of lion hunting behaviour. In simplified form, the model is equations (1) and (2) with Z_t a random variable given by

$$\text{probability } (Z_t = E/n) = \lambda_n, \quad \text{probability } (Z_t = 0) = 1 - \lambda_n$$

where X_t denotes stomach contents at the beginning of day t , Z_t is food intake, n denotes hunting group size, α is the daily minimum food requirement per lion ($= 6$ kg), Cap is the stomach capacity ($= 30$ kg) and E is the average food content per prey item ($= 164$ kg for a zebra and 12 kg for a gazelle; zebra carcasses last up to three days; gazelle are consumed immediately). The kill probabilities λ_n depend on group size n , and also on the type of prey and habitat. We wish to find the behaviour that maximizes survival over a time period of T days. Thus equation (3) holds, and the resulting dynamic programming equation is similar to that given by equations (4)–(6) above but with $\beta_i = 0$, because lions are not preyed upon.

Typical predictions of this model are shown in Table 1a, which shows survival probabilities P and optimal hunting group sizes n^* as functions of initial stomach contents x ($T = 30$ days), for three sets of prey/habitat combinations.

Notice that the optimal group size n^* predicted by the dynamic model depends on the lion's current state x , as well as on prey type and habitat. For example, well-fed lions should hunt zebra in groups of size six, rather than two as predicted by Caraco and Wolf's model. In fact, hunting in groups of size two is severely suboptimal, yielding only a 78% probability of survival over a 30-day period. The main reasons for this result are (i) small groups hunting zebra have reduced kill probabilities ($\lambda_2 = 0.30$, $\lambda_6 = 0.43$), and (ii) two lions cannot consume a whole zebra before the carcass rots so that much meat is wasted if only two lions share a zebra kill. The importance of this constraint

Table 1 Probability of survival and hunting group size

x (kg)	a						b		
	(1)		(2)		(3)		N	P	n^*
	P	n^*	P^a	n^*	P	n^*			
5	0.86	3	0.26	2	0.06	1	2	0.18	2
10	0.90	4	0.43	2	0.13	1	4	0.75	2
20	0.97	4	0.64	2	0.26	2	6	0.95	2
30	0.99	6	0.74	2	0.35	2	8	0.993	2
							10	0.999	2

a , Maximum probability of survival (P) and optimal hunting group size (n^*), as functions of current stomach contents (x) with a 30-day horizon, for lions hunting (1) zebra in wet-season Serengeti habitat, (2) Thomson's gazelle in wet-season habitat, and (3) gazelle in dry-season habitat. From Clark²¹. b , Maximum probability of survival (P) and optimal hunting group size (n^*) for prides that communally share kills, as a function of pride size (number of adult females), N . Prey is zebra, dry-season habitat.

is made obvious by the dynamic approach, but was overlooked in previous analyses of lion hunting behaviour.

A second prediction of the dynamic model is that hungry lions should hunt in smaller groups than well-fed lions. In times of prey scarcity, for example, lions may be hungry most of the time, and our prediction is that smaller hunting groups will then be observed. The breakup of lion prides during the dry season in the Kalahari has been described by Owens and Owens²².

This model can be modified to allow for scavenging of kills by pride mates, which Packer suggests as the main mechanism underlying lion social behaviour. Consider a pride containing N adult females, and suppose that the pride hunts in k separate groups each containing n females so that $N = kn$. Kills are assumed to be shared equally among pride members. These assumptions lead to a binomial distribution for Z_t .

Table 1b shows the optimal hunting group sizes n^* and 30-day survival probabilities P for this model of communal scavenging, as a function of pride size N . The corresponding results for a pride of size $N = 6$ that does not share kills are $n^* = 6$ and $P = 0.42$, indicating that communal scavenging of kills may greatly increase survival rates for individual lions under certain circumstances—namely where prey is large, relatively scarce, and perishable, and where ecological conditions permit separate hunting groups to communicate to pride mates the fact that a kill has been made. These circumstances are characteristic of the African savannah.

Small birds in winter

During the winter small birds must forage most of the daylight period if they are to obtain enough food to satisfy their daily requirement²³. Under these conditions adverse fluctuations in the amount of food obtained can easily result in starvation. Foraging activities also expose a bird to predators, so that there is often a trade off between the need to obtain food and the need to avoid predators.

To model the foraging decisions of a bird over a single day, let time $t = 0$ be dawn and let time $t = T$ be dusk. At each of the times $t = 0, 1, \dots, T-1$ the bird must choose a foraging option from some set of available options. The option chosen at time t determines the distribution of the amount of food obtained between time t and $t+1$ and the predation risk. The option also determines metabolic consumption between t and $t+1$. Overnight the bird uses x_c units of energy reserves.

To model foraging behaviour in winter, we assume that each day in winter can be described in this way. Given a terminal reward (expected future reproductive success) for the various energy states at the end of winter, the optimal policy on each day can be found by working backwards from the terminal reward. When there are sufficiently many days left in winter,

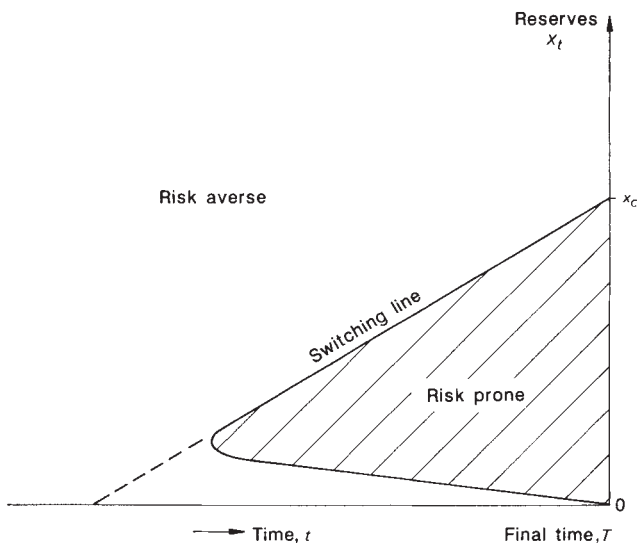


Fig. 1 The region of risk-prone behaviour as a function of energy reserves and time of day. An animal has two possible courses of action, a_H and a_L , of high and low variance, respectively, in the energetic gain resulting from the course of action. The dashed line has slope μ equal to the mean rate of gain under both foraging options. In the shaded region of the figure it is optimal to choose the 'risky' option a_H .

the optimal policy on all days is the same. This policy is the one that maximizes the bird's long-term survival probability.

This sort of model can be used to analyse the value of a given increase in energy reserves, the trade off between starvation and predation risks, or the importance of variability in foraging returns. We discuss some results concerning the last of these topics.

The calculations that yield the optimal policy also give the long-term survival probability as a function of energy reserves. In particular, the survival probability of a bird at dusk is a non-linear function of its energy reserves. This function is convex for low levels of reserves and concave for high levels of reserves. One can regard this function as a terminal reward function at dusk. The consequences of non-linear reward functions are well-known (see ref. 12 for review). Given a choice between two options that result in the same mean gain but have different variances, it is optimal to maximize variability (risk-prone behaviour) when the derivative is increasing and to minimize variability (risk-averse behaviour) when the derivative is decreasing. In the small-bird-in-winter paradigm the situation is complicated by the need to avoid daytime starvation, and can be summarized as follows.

Suppose that two actions a_H and a_L are available at each of the times of day $t = 0, 1, \dots, T-1$. These actions have no predation risk and result in the same mean energetic gain μ per time unit, but the variance in gain is greater under a_H than under a_L . We refer to a_H as the high-variance option and a_L as the low-variance option. The general form of the optimal policy is shown in Fig. 1. The switching line in this figure has slope μ . An animal whose reserves are below this line will not survive the night if it obtains energy at the mean rate. It is thus optimal to take risks in the shaded region of the graph by choosing the high variance option a_H . When reserves are very low, it is always optimal to be risk-averse and choose the low variance option, because the high variance option has a greater risk of immediate starvation. An animal with reserves above the switching line expects to survive the night and therefore plays safe by being risk-averse. The result of these effects is to give a wedge-shaped region in which it is optimal to be risk-prone.

Given a choice between several actions we define the canonical cost of choosing a particular action as the loss in expected reproductive success which is incurred as a result of this choice¹⁰. In the small-bird-in-winter model the canonical cost of action a , $c(a, x, t)$, depends on reserves and time of day, and is essentially the reduction in overwinter survival probability which results from choosing a .

In this example the canonical cost of choosing a_L when a_H is optimal and of choosing a_H when a_L is optimal can be quite high. But the probability of being in the shaded region in which a_H is optimal is small. Houston and McNamara¹⁷ use this to argue that the selection pressure to be risk-prone under the appropriate circumstances is much smaller than the corresponding selection pressure to be risk-averse. This conclusion is supported by analysing the total mortality under various strategies. If a bird with a single behavioural option is given a further option yielding the same mean amount of food but having a higher variance, the bird can slightly reduce its probability of starvation by choosing the high variance option in appropriate circumstances. If instead the second option has a lower variance it is optimal for the bird to use this option most of the time, and the resultant drop in starvation probability is usually massive. In line with this, the data suggest that risk-prone behaviour is a much less robust phenomenon than risk-averse behaviour (see refs 17, 24 for further discussion and references).

The dawn chorus

In spring, many songbirds produce an intense burst of song at dawn, the dawn chorus. McNamara *et al.*²⁵ use a dynamic model to analyse the factors that can produce this chorus. The model considers a male bird that sings to attract a mate. A period of five days is modelled. The daylight period of each day is divided into 96 equal time intervals. During each interval the bird can either forage or sing. If it forages it gains no mate but finds food of energy value Z_t . The energy expended during the interval is a linearly increasing function of the bird's energy reserves (and hence body mass). If reserves at time t are x then reserves at time $t+1$ satisfy

$$X_{t+1} = x + Z_t - (a_F + k_F x)$$

If the bird sings it finds no food and its reserves at time $t+1$ are given by

$$X_{t+1} = x - (a_S + k_S x)$$

A singing bird attracts a mate with probability m_t . A bird which manages to attract a mate can spend the remaining time foraging. Birds rest between dusk and the following dawn. Overnight energy consumption is assumed to follow a normal distribution with mean μ_N and variance σ_N^2 .

During the five-day period a bird starves if its reserves fall to zero. If the bird survives till dawn on the sixth day it receives a reward of one unit. If it has also obtained a mate by this time it receives a further reward. The optimal policy maximizes

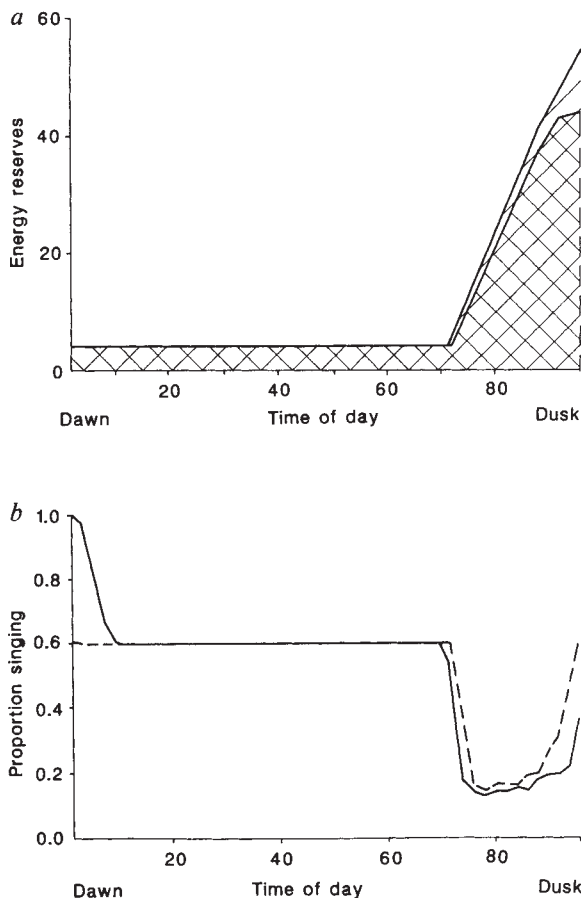


Fig. 2 *a*, The optimal policy in the model of singing and foraging. The bird can either sing or forage. When there is a constant loss of energy each night ($\mu_N > 0$, $\sigma_N^2 = 0$) it is optimal to forage if and only if reserves are in the double hatched region. When $\sigma_N^2 = 9$ it is optimal to forage if and only if reserves are in the single or double hatched region. *b*, The proportion of birds singing under the two policies shown in *a*. When $\sigma_N^2 = 0$, the proportion is given by the broken line and there is no dawn chorus. When $\sigma_N^2 = 9$, the proportion is given by the solid line. There is a marked dawn chorus and a slight reduction in singing at dusk.

expected total reward. In general the optimal action depends on reserves and time of day, but is a very similar function of these on all days except the last.

In describing the results of the model, it is important to distinguish between the optimal policy and the expected behaviour that is generated by following the optimal policy. The optimal policy is a specification of a critical level of reserves $\hat{x}_t$, such that if $X_t > \hat{x}_t$ then it is optimal to sing and if $X_t < \hat{x}_t$ then it is optimal to forage. The expected behaviour depends on the proportion of time spent above or below $\hat{x}_t$ and gives the proportion of birds singing in a large population of identical birds. In the absence of an energy loss overnight, $\hat{x}_t$ is constant and there is no circadian rhythm. Adding a constant loss of energy per night (that is, $\mu_N > 0$, $\sigma_N^2 = 0$) results in a rise in $\hat{x}_t$ in the afternoon as a bird must build up its reserves to survive the following night (Fig. 2*a*). This results in a dip in the proportion of birds that are singing in the afternoon (Fig. 2*b*). As might be expected, a burst of song at dawn can be produced by increasing the probability of getting a mate m_t at this time or decreasing the returns from foraging Z_t . McNamara *et al.*²⁵ found, however, that making the energy lost overnight a random variable ($\sigma_N^2 > 0$) could also produce a dawn chorus. The reason is that a bird must get its reserves at dusk above the amount

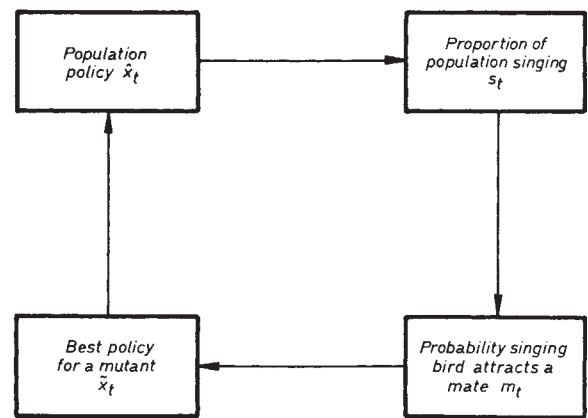


Fig. 3 The iterative procedure for finding a dynamic ESS (after Houston and McNamara²⁸). The procedure starts with a population following the policy 'sing if reserves at time t are greater than $\hat{x}_t$.' From this policy, the proportion of the population singing at time t , s_t , is calculated. Using the interaction between singing males, the probability m_t that a singing bird attracts a mate at time t is found. From m_t , the best policy $\hat{x}_t$ for a single mutant can be found. The population policy is then replaced by $\hat{x}_t$ and the process is repeated.

required to survive an unusually bad night. To do so, it sacrifices some singing at dusk (Fig. 2). On most nights it does not use up all the reserves that it has accumulated, and so starts the day with its reserves above $\hat{x}_t$. It sings until its reserves fall to $\hat{x}_t$. In this case the bird sings at dawn not because dawn is a good time for singing but because the problems posed by variable energy loss overnight require it to trade some singing at dusk for some singing at dawn. Thus the intensity of the dawn chorus may depend on the fluctuations in overnight temperature. Reid found a relationship between overnight temperature and song in the Ipswich sparrow²⁶.

This example can be used to introduce the idea of dynamic evolutionarily stable strategies. An evolutionarily stable strategy (ESS) renders a population safe from invasion by a mutant that adopts a different strategy²⁷. Most previous ESS models have been concerned with a single choice of action that does not depend on the animal's state. If the probability of attracting a mate by singing depends on whether or not other birds are singing, then a game-theory analysis is required. A dynamic ESS is a rule for determining a sequence of actions as a function of an animal's state, the behaviour of other birds, and time of day. Houston and McNamara²⁸ consider a large population limit in which the probability of attracting a mate depends on the proportion of other birds that are singing. The procedure for finding the ESS, as shown in Fig. 3, is based on finding the best response for a mutant when a population follows a given strategy. For an initial population strategy, the best response for a mutant is found, and the population strategy is then replaced by the best mutant strategy. The procedure is repeated until the best response for a mutant is the current population strategy. This strategy is then the ESS for the population. The procedure can be applied to a wide range of games; Houston and McNamara²⁹ use it to analyse a dynamic version of the hawk-dove game.

Clutch size in parasitic insects

Parasitic insects, such as the wasp *Nasonia vitripennis* or various tephritid fruit flies, provide an ideal setting for the study of reproductive behaviour using dynamic models. The adult of these species is free living, laying its eggs in the host. We begin by identifying one or more state variables. Natural choices are energy reserves, mature eggs remaining, or a combination of

mature eggs and oocytes (potential eggs). For simplicity, we will use only the number of mature eggs remaining as a state variable. The 'reward function' for parasitic insects is related to the fecundity of mature offspring from a particular clutch³⁰. This function involves a combination of survivorship, size, and egg production of offspring. Also of particular importance for parasitic insects is adult survivorship from one potential host to another (or from one day to another).

We consider the behaviour of an insect that starts its search for hosts with a complement of X_0 mature eggs, and let X_t denote the number of eggs remaining at the start of period t . We assume that foraging for hosts stops at some period T . For example, if eggs are resorbed at the end of a day, then T would denote the number of search periods in a day. Alternatively, if we are interested in a lifetime problem, T might be the time at which the insect dies or the time at which its eggs no longer hatch. We then define $J(x, t, T)$ as the maximum expected fitness from laying eggs between t and T , given that $X_t = x$.

We are interested in computing $J(X_0, 1, T)$. Assume that there are n different kinds of hosts (indexed by volume, fruit or insect type) and that λ_i is the probability of encountering a host type i in a period of unit length. If a clutch of size C is laid in a host of type i , the fitness accrued to the mother is $f_i(C)$ and the handling time is $\tau_i(C)$. We specify survival through a function $P_i(x, t)$ representing the probability of surviving from period t to period $t+1$, given that the insect is alive at period t , that $X_t = x$, and that a host of type i is encountered. The index $i=0$ is used to indicate that no host is encountered in a given period, $\lambda_0 = 1 - \sum \lambda_i$ is the probability of no encounter, and $P_0(x, t)$ is the survival probability. With these assumptions, the fundamental equation is

$$J(x, t, T) = \sum_{i=0}^n \lambda_i \max_{C \leq x} (f_i(C) + P_i(x, t) \times J(x - C, t + \tau_i(C), T)) \quad (7)$$

with $f_0(C) = 0$ for all values of C , and $t + \tau_i(C)$ is replaced by T if it exceeds T . This equation shows that the interplay of time and state variable is crucial for the analysis of insect oviposition decisions.

Mangel¹⁶ analyses a simple version of equation (7) in which all $\tau_i(C) = 1$ and $P_i(x, t) = P_0 = \text{constant}$. This model can explain observed frequency distributions of clutch sizes³⁰, some aspects of superparasitism³¹, and the response of the apple maggot *Rhagoletis pomonella* (Walsh) to its oviposition-marking pheromone³¹. In the first case, small clutches predominate in

the observed frequency distribution of clutches of the parasitic wasp *N. vitripennis*. This is exactly what is predicted by equation (7). The frequency of small clutches should increase with the time left for searching and the probability of encountering a host. The model can also be used to predict when an insect should oviposit in a previously parasitized host. If the remaining search time is great or the insect has few eggs, then a previously parasitized host should be rejected. As t approaches the time horizon T , the theory shows that hosts which would be considered unacceptable for small values of t become acceptable. In all three cases, the interaction of the time horizon and state variable is needed for understanding the oviposition problem. We see this as a general feature of dynamic models.

Conclusion

The technique for analysing behaviour we present here has the following advantages: (i) It takes account of the state of the animal and how that state changes according to the animal's actions and the environment; (ii) it provides a common currency for assessing behavioural choices in terms of overall fitness, which can be used to analyse trade-offs between different actions; (iii) it includes constraints on state variables or behaviour. Models based on this technique often include a higher degree of biological realism and lead to predictions and insights not provided by simpler models. Stochastic dynamic programming provides a method for finding the optimal behaviour (or behaviours) within the dynamic framework. As a computational technique, stochastic dynamic programming has certain limitations, perhaps the most serious of which is the 'curse of dimensionality' in which computational needs grow vastly as the number of state variables increases. Hence, there is an upper limit to the degree of biological complexity that the method can realistically encompass.

We have applied this dynamic approach to many other situations, such as diel vertical migration of aquatic organisms³², optimal choice of prey items³³, diving behaviour of water birds, growth and migration of salmon, sex change in slugs, web locations of spiders, and food hoarding by small birds.

In our main examples, there are no interactions between animals. We have mentioned that under some circumstances evolutionarily stable dynamic strategies can be modelled in a relatively simple way. In general, the development of such models is likely to encounter conceptual and computational difficulties. Nevertheless, we believe that this is a biologically important issue that deserves further theoretical and empirical research, and that the technique of stochastic dynamic programming will be a powerful tool in this enterprise.

1. Darwin, C. *On the origin of species by natural selection* (Murray, London, 1859).
2. Oster, G. F. & Wilson, E. O. *Caste and Ecology in the Social Insects* (Princeton University Press, 1978).
3. Maynard Smith, J. *A. Rev. Ecol. Syst.* **9**, 31-56 (1978).
4. McFarland, D. J. *Nature* **269**, 15-21 (1977).
5. Stearns, S. C. *Q. Rev. Biol.* **51**, 3-47 (1976).
6. Schaffer, W. M. *Am. Nat.* **121**, 418-431 (1983).
7. Sibly, R. & Calow, P. *J. theor. Biol.* **102**, 527-547 (1983).
8. Tinbergen, N. *et al. Behaviour* **19**, 74-117 (1962).
9. Mangel, M. & Clark, C. W. *Ecology* **67**, 1127-1138 (1986).
10. McNamara, J. M. & Houston, A. I. *Am. Nat.* **127**, 358-378 (1986).
11. Schoener, T. A. *Rev. Syst. Ecol.* **2**, 369-404 (1971).
12. Stephens, D. W. & Krebs, J. R. *Foraging Theory* (Princeton University Press, 1986).
13. Werner, E. E. & Gilliam, J. F. *A. Rev. Ecol. Syst.* **15**, 393-425 (1984).
14. Ross, S. *Introduction to Stochastic Dynamic Programming* (Academic, New York, 1983).
15. McCleery, R. M. *Anim. Behav.* **25**, 1005-1015 (1977).
16. Mangel, M. *J. Math. Biol.* **25**, 1-22 (1987).
17. Houston, A. I. & McNamara, J. M. in *Relevance of models and theories in ethology* (eds Campan, R. & Zayan, R.) 61-75 (Privat, I.E.C., Toulouse).
18. Packer, C. in *Ecological Aspects of Social Evolution* (eds Rubenstein, D. I. & Wrangham, R. W.) 429-451 (Princeton University Press, Princeton, 1986).
19. Caraco, T. & Wolf, L. L. *Am. Nat.* **109**, 343-352 (1975).
20. Schaller, G. B. *The Serengeti Lion* (University of Chicago Press, Chicago, 1972).
21. Clark, C. W. *Anim. Behav.* **35**, 361-368 (1987).
22. Owens, M. J. & Owens, D. D. *Int. Wildl.* **14**, 4-13.
23. Gibb, J. A. *Ibis* **96**, 513-543 (1954).
24. McNamara, J. M. & Houston, A. I. *Acta Biotheoretica* **36**, 3-22 (1987).
25. McNamara, J. M., Mace, R. H. & Houston, A. I. *Behav. Ecol. Sociobiol.* **20**, 399-405 (1987).
26. Reid, M. L. *Anim. Behav.* **35**, 1735-1743 (1987).
27. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, 1982).
28. Houston, A. I. & McNamara, J. M. *J. theor. Biol.* **129**, 57-68 (1987).
29. Houston, A. I. & McNamara, J. M. *Evol. Ecol.* (in the press).
30. Charnov, E. L. & Skinner, S. W. *Flor. Ent.* **67**, 5-21 (1984).
31. Roitberg, B. D. & Prokopy, R. J. *Ecol. Ent.* **7**, 437-442 (1982).
32. Clark, C. W. & Levy, D. A. *Am. Nat.* (in the press).
33. Houston, A. I. & McNamara, J. M. *Behav. Ecol. Sociobiol.* **17**, 135-141 (1985).

Self-tolerance eliminates T cells specific for Mls-modified products of the major histocompatibility complex

John W. Kappler^{*†‡}, Uwe Staerz^{||}, Janice White^{*} & Philippa C. Marrack^{*†§}

^{*} Howard Hughes Medical Institute and Division of Basic Immunology, Department of Medicine, National Jewish Center for Immunology and Respiratory Medicine, 1400 Jackson Street, Denver, Colorado 80206, USA

[†] Department of Microbiology and Immunology, [‡] Department of Medicine, [§] Department of Biochemistry, Biophysics and Genetics, University of Colorado Health Sciences Center, Denver, Colorado 80206, USA

^{||} Basel Institute for Immunology, Basel, Switzerland

In mice the product of the Mls^a locus is an unusual antigen capable of interaction with certain products of the major histocompatibility locus (MHC) to form a ligand for a large portion of the T-cell α/β receptor repertoire, including nearly all receptors that use V _{β 8.1}. The presence of Mls^a/MHC during T-cell development results in the deletion of T cells that express V _{β 8.1}, documenting the importance of clonal deletion in establishing tolerance to self antigens.

THE T-cell $\alpha\beta$ receptor generally recognizes ligands composed of a fragment of a foreign antigen complexed to a product of the major histocompatibility complex (MHC)^{1,2}. Polymorphic amino-acid residues within the MHC molecule control both the binding of antigen fragments and the interaction of the antigen/MHC complex with the T-cell $\alpha\beta$ receptor, suggesting that the receptor contacts both the antigen and MHC portion of the ligand. This phenomenon is often referred to as MHC-restricted antigen presentation^{3,4}. The recent report of the three-dimensional structure of a class I human MHC molecule⁵ has provided a potential molecular explanation for MHC restriction. The molecule contains a well-defined cleft as the probable binding site for peptide antigens. When this site is occupied, the structure predicts that residues of the antigenic peptide and a number of the polymorphic residues of the MHC molecule face outward⁶ offering a large area for potential interaction with the T-cell $\alpha\beta$ receptor.

T-cell $\alpha\beta$ receptors are involved in two important events during T-cell development which shape the repertoire of T-cell receptors available for antigen/MHC recognition. First, in the thymus, interactions between the T-cell $\alpha\beta$ receptor and MHC molecules positively select a T-cell repertoire skewed towards recognition of antigens complexed with self rather than with foreign MHC molecules⁷⁻¹¹. Which, if any, self-antigens are bound to the MHC molecules during this selection is unknown. Second, in a process called self-tolerance, T cells in this selected pool which can react with self-MHC/self-antigens are inactivated¹². The mechanisms of these two processes are unknown and this paradoxical dual role of the interaction between $\alpha\beta$ receptors and self-MHC in positive and in negative selection of the T-cell repertoire remains poorly understood.

The frequency of T cells responsive to most antigen/MHC ligands is quite low, reflecting perhaps the complex interaction between the variable portions of the $\alpha\beta$ receptor and the antigen/MHC ligand. But there are some exceptions to this general rule. In mouse the products of the Mls locus are a well-documented example. A very high proportion of T cells from Mls^b mice can respond to leukocytes expressing the Mls^a allele¹³⁻¹⁸. This suggests that Mls^a may be an unusual antigen, which when combined with certain products of the murine MHC, H-2, creates a 'degenerate' or 'permissive' ligand that can interact with a high proportion of T-cell $\alpha\beta$ receptors and, therefore, may be a useful model system for antigen presentation. Another view, however, is that a second receptor less diverse than the $\alpha\beta$ receptor controls Mls^a recognition^{19,20}.

Here we present our studies on the repertoire of T cells whose $\alpha\beta$ receptors use members of the V _{β 8} family. We have found a

close linkage between the use by T cells of V _{β 8.1} and their reactivity to Mls^a in combination with products of some, but not all, H-2 haplotypes, establishing the involvement of the $\alpha\beta$ receptor and MHC in Mls^a recognition. We found that mice expressing both Mls^a and the appropriate H-2 products eliminated a large proportion of V _{β 8.1} T cells during T-cell development in the thymus. These results document elimination of self-reactive clones as one of the mechanisms of self-tolerance to self-antigen/self-MHC complexes. In addition, they point out that the ability of the Mls^a/MHC ligand to be recognized by such a high proportion of T cells makes it particularly useful in studies of the role of receptor/ligand interactions during T-cell development and perhaps in the response to more conventional antigens.

Monoclonal antibodies

Three monoclonal antibodies have been reported which detect members of the V _{β 8} family. F23.1 binds to T-cell $\alpha\beta$ receptors bearing any of the three members of this family^{21,22}. KJ16 binds to receptors bearing either V _{β 8.1} or V _{β 8.2}, but not V _{β 8.3} (refs 22, 23). A third antibody²⁴, F23.2, resulting from the immunization that also produced F23.1 was shown to bind a subset of KJ16⁺ receptors, but its exact specificity was not known. We established that this antibody was specific for V _{β 8.2} by examining V _{β} usage in a set of 20 KJ16⁺ T-cell hybridomas and two KJ16⁺ T-cell clones. These T cells were of various antigen/H-2 specificities and from a variety of strains. Our analysis is summarized in Table 1. We determined which of the KJ16⁺ clones or hybridomas bound F23.2 by cytofluorometric analysis. Fluorescence histograms are shown for two of the hybridomas in Fig. 1a. Both T-cell clones and 12 of the 20 T-cell hybridomas were F23.2⁺. We then used a number of molecular biological techniques to determine whether the T cells were using V _{β 8.1} or V _{β 8.2} and, for the hybridomas, to which J _{β} element this V _{β} segment had rearranged. Immediately apparent from these data was the perfect correlation between F23.2 binding and V _{β 8.2} expression. Furthermore, the data confirm that all KJ16⁺ T cells used either V _{β 8.1} or V _{β 8.2} and that J _{β} usage was not a factor in the specificity. Therefore, in subsequent studies KJ16 and F23.2 were used to distinguish T cells using V _{β 8.1} (KJ16⁺, F23.2⁻) from those using V _{β 8.2} (KJ16⁺, F23.2⁺).

Mls-reactive T cells

We examined the specificity of T-cell $\alpha\beta$ receptors containing either V _{β 8.1} or V _{β 8.2}. In general the reactivity of an $\alpha\beta$ receptor is dependent on the sum of its variable elements, with no single component defining the specificity. But we have described an

Table 1 Monoclonal antibodies that distinguish $V_{\beta 8.1}^+$ and $V_{\beta 8.2}^+$ T-cell receptors

T-cell clone hybridoma	Strain of origin	Specificity	Monoclonal antibody binding		$V_{\beta 8}$ usage (1 or 2)	DNA seq.	Method used for $V_{\beta 8}$ assignment		J_{β} usage	Reference
			KJ16	F23.2			Del. map.	Re-arrange.		
2C	BALB.B	H-2D ^d	+	+	2	+			?	41
D10	AKR/J	ConAlb+I-A ^k	+	+	2	+			?	42
DO-11.10	BALB/c	cOVA+I-A ^d	+	+	2	+		+	1.1	43, 44
21DO-90.3	BALB/c	cOVA+I-A ^d	+	+	2			+	1.1	
3BBM-78	C57BL/10	I-A ^{bm12}	+	+	2			+	1.1	
17BBM52	C57BL/10	I-A ^{bm12}	+	+	2		+	+	1.2	
BO-97.10	BALB.B	cOVA+I-A ^b	+	+	2		+		1.2	
17BBM-20	C57BL/10	I-A ^{bm12}	+	+	2			+	1.4	
21DO-18.6	BALB/c	cOVA+I-A ^d	+	+	2		+	+	2.4	
18'BBM4	C57BL/10	I-A ^{bm12}	+	+	2			+	2.4	
21DO-1.1	BALB/c	cOVA+I-A ^d	+	+	2			+	2.5	
17BBM87	C57BL/10	I-A ^{bm12}	+	+	2		+	+	2.6	
SKK-45.10	(C58×SJA)	KLH+I-A ^k	+	+	2		+	+	2.6	23
18BBM4	C57BL/10	I-A ^{bm12}	+	+	2			+	2.6	
3BBM-74	C57BL/10	I-A ^{bm12}	+	—	1		+	+	1.1	
3BBM-51	C57BL/10	I-A ^{bm12}	+	—	1			+	1.3	
BO-12.27	BALB.B	cOVA+I-A ^b	+	—	1			+	1.4	
3DT-52.5	BALB/c	H-2D ^d	+	—	1	+		+	2.3	44, 45
8DO-51.15	BALB/c	cOVA+I-A ^d	+	—	1		+	+	2.3	
8DO-119.7	BALB/c	cOVA+I-A ^d	+	—	1		+		2.5	
D1G10G11	C57BL/6	I-E ^d	+	—	1			+	2.6	23
18BBM67	C57BL/10	I-A ^{bm12}	+	—	1			+	2.6	

T-cell clones 2C and D10 were obtained from Dr Herman Eisen and Dr Charles Janeway respectively. The rest of the T cells were hybridomas produced and characterized as previously described^{39,40}. References are given for lines previously reported. Binding of KJ16 and F23.2 was determined as in Fig. 1. Three methods were used to determine $V_{\beta 8.1}$ or $V_{\beta 8.2}$ usage. First, in four cases the functional $V_{\beta 8}$ gene had been cloned and sequenced. Data for 2C were obtained from Dr Susumu Tonegawa, and for D10 from Dr Janeway. The sequences for DO-11.10 and 3DT-52.5 were previously reported⁴⁴. Second, in several cases we determined $V_{\beta 8}$ usage by deletion mapping. The $V_{\beta 8.2}$ element is upstream from $V_{\beta 8.1}$ in the germline⁴⁶. Therefore, the rearrangement of $V_{\beta 8.2}$ to the J_{β} cluster deletes $V_{\beta 8.1}$ on the functional chromosome, whereas rearrangement of $V_{\beta 8.1}$ leaves $V_{\beta 8.2}$ undisturbed. In those KJ16⁺ T cells where a non-functional rearrangement on the second chromosome had deleted both $V_{\beta 8.1}$ and $V_{\beta 8.2}$, a $V_{\beta 8.2}$ gene in germline configuration determined by Southern blot analysis was diagnostic for those T-cell lines using $V_{\beta 8.1}$ rather than $V_{\beta 8.2}$. We are able to perform this type of analysis with eight of the T-cell hybridomas. Finally, the pattern of $V_{\beta 8}$ rearrangement in 18 hybridomas was analysed by Southern blot. Restriction enzymes were chosen to give widely different fragment sizes for rearrangement of $V_{\beta 8.1}$ as opposed to $V_{\beta 8.2}$ to each of the J_{β} clusters. A $V_{\beta 8}$ probe was used that detects all three members of the family. *HindIII* was used to analyse rearrangements to the $J_{\beta 1}$ cluster and *TaqI* for rearrangements to the $J_{\beta 2}$ cluster. In the first case $V_{\beta 8.1}$ rearrangements yielded DNA fragments in the range of 5.9 to 7.5 kilobases (kb), depending which J_{β} element in the cluster was involved, whereas $V_{\beta 8.2}$ rearrangements gave fragments of 14.3–16.0 kb. The corresponding range in the second case was 0.9–2.0 kb for $V_{\beta 8.1}$ and 9.7–10.8 kb for $V_{\beta 8.2}$. J_{β} usage was determined by Southern blot analysis using the enzymes *HindIII*, *TaqI* and *EcoRI* and probes specific for the $J_{\beta 1}$ or $J_{\beta 2}$ cluster.

exception in a previous study^{25,26} in which nearly all $\alpha\beta$ receptors using $V_{\beta 17a}$ were found to bind to ligands involving some or all of the allelic forms of the class II H-2 molecule, I-E. We looked for a similar predominant specificity for $V_{\beta 8.1}^+$ or $V_{\beta 8.2}^+$ receptors, including reactivity to Mls in our screen, because we had previously found a possible link between Mls type and $V_{\beta 8}$ usage²⁷. Our analysis was simplified by the recent results indicating that the genetics of the Mls locus are not as complex as once thought^{28,29}. The data are now consistent with two alleles of this locus, *a* and *b*. The highly stimulatory determinant for T cells is produced in Mls^a mice only, so that T cells from Mls^b mice respond strongly to leukocytes from Mls^a mice, but not vice versa.

Therefore, we isolated KJ16⁺ T cells from B10.BR (Mls^b, H-2^k) mice, used them to produce a random set of KJ16⁺ T-cell hybridomas, divided the hybridomas into $V_{\beta 8.1}^+$ and $V_{\beta 8.2}^+$ using F23.2 staining, and analysed the reactivity of each set to spleen cells from CBA/J and AKR/J (Mls^a, H-2^k) mice. The reactivity was detected by measuring the stimulation of IL-2 production by the T-cell hybridomas. The results are summarized in Table 2. Of the hybrids, approximately one third used $V_{\beta 8.1}$ and the rest, $V_{\beta 8.2}$. Four of thirty (13%) $V_{\beta 8.2}$ hybrids tested reacted to Mls^a, the approximate proportion seen among randomly-tested T-cell hybrids using any of a variety of V_{β} elements, reflecting the high frequency of Mls^a reactivity. Even so, the frequency of Mls^a-reactive cells was even higher among $V_{\beta 8.1}$ hybrids. Of 19 hybrids tested, 15 (~80%) reacted to Mls^a. These results

clearly establish the linkage between the $\alpha\beta$ receptor and recognition of Mls^a and indicate that $V_{\beta 8.1}$ probably makes a major contribution to the ability of $\alpha\beta$ receptors to react to Mls^a, dominating contributions from other variable parts of the receptor.

The involvement of the $\alpha\beta$ receptor in Mls^a recognition indicates that Mls^a should associate with H-2 products to provide a ligand for these receptors. Because this type of recognition usually requires a product of self H-2 as part of the ligand, we tested the four $V_{\beta 8.2}^+$ and 12 of the 15 $V_{\beta 8.1}^+$ Mls^a-reactive hybrids for response to Mls^a presented by spleen cells of H-2 types other than H-2^k on the DBA/1(D1) or DBA/2(D2) background (Table 3). In addition to a response to Mls^a presented by AKR/J(H-2^k) and CBA/J(H-2^k) cells, all of the $V_{\beta 8.1}^+$ hybrids reacted to Mls^a on DBA/2(H-2^d) and D1.LP(H-2^b) cells, although in some cases either H-2^d or H-2^b cells presented Mls^a rather poorly compared to H-2^k cells. No $V_{\beta 8.1}$ hybrids responded to Mls^a on DBA/1(H-2^q) or D2.GD (recombinant H-2^d/H-2^b) cells. The four $V_{\beta 8.2}^+$ Mls^a-reactive hybrids had a distinctly different H-2 specificity. These hybrids failed to react to Mls^a presented by any cells other than H-2^k. These results indicate that, as with conventional antigens, Mls^a is recognized by the $\alpha\beta$ receptors in combination with H-2 products, except that there is not always a strict requirement for self H-2. These results also indicate that the V_{β} composition of the $\alpha\beta$ receptor strongly influences which combinations of Mls^a and H-2 can be recognized. Interestingly, others^{17,18} have reported strong

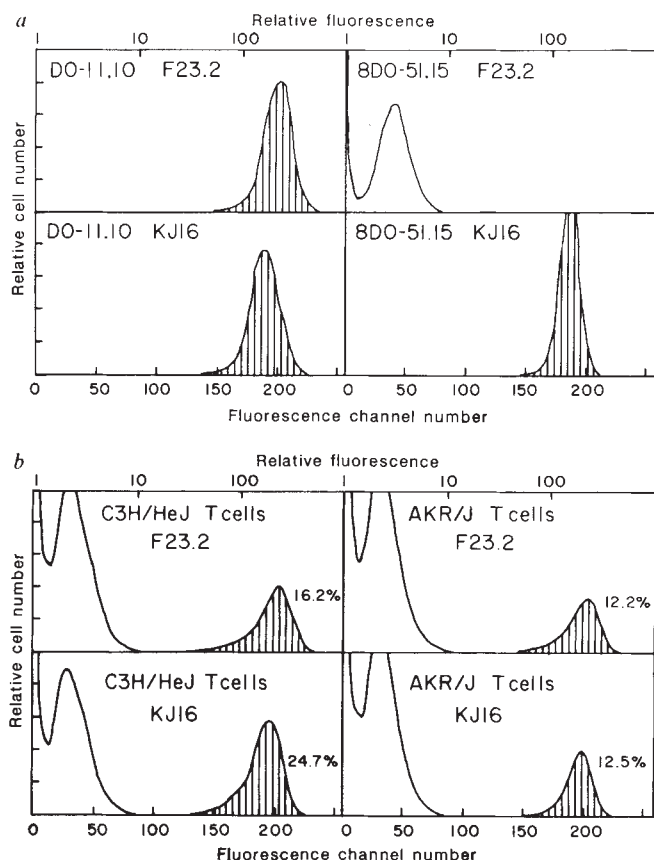


Fig.1 Binding of KJ16 and F23.2 to T-cell hybridomas and lymph node T cells. *a*, T-cell hybridomas DO-11.10 and 8DO-51.15 were analysed for binding of KJ16 and F23.2 using biotinylated forms of the antibodies with phycoerythrin-streptavidin (Becton-Dickinson). Analyses were performed with an Epics C Flow Cytometer as previously described²⁵. Each histogram shows the analysis of 5,000 cells. Hatched areas, cells staining significantly brighter than the control cells (no biotinylated antibodies). *b*, Nylon fibre purified T cells from C3H/HeJ (Mls^b, H-2^k) and AKR/J (Mls^a, H-2^k) lymph nodes were stained with KJ16 and F23.2 as for *a*. Fifty thousand cells were analysed for each histogram. Hatched areas, as in *a*. The percent of cells staining with each antibody is shown.

T-cell responses to Mls^a presented on H-2^k and H-2^d cells, but more variable results with H-2^b and the recombinant H-2^d/H-2^b haplotype of D2.GD and virtually no response with H-2^a. Given the relative lack of mouse strains carrying recombinant H-2 haplotypes and Mls^a, we cannot formally map the relevant MHC products involved in Mls^a recognition; however, the response to DBA/2 (I-A^d/I-E^d) cells versus the lack of response to D2.GD (I-A^d only) cells would suggest a possible role of the I-E molecule, at least in the H-2^d haplotype.

Suppression of V_{β8.1} expression

The association between V_{β8.1} and Mls^a/H-2 recognition raised the question of the fate of V_{β8.1} T cells in Mls^a animals. Our previous studies with V_{β17a} indicated that, correlating with the strong tendency of V_{β17a}-containing αβ receptors to recognize I-E ligands, a large proportion of V_{β17a} T cells were deleted in I-E⁺ mice at an early stage in development²⁶. Therefore, we examined a number of mouse strains differing in H-2 and Mls type using KJ16 and F23.2 to determine the relative expression of V_{β8.1} and V_{β8.2} on peripheral T cells. Purified lymph-node T cells were stained with KJ16 or F23.2 and the proportion of cells expressing V_{β8.1} and V_{β8.2} calculated. Sample fluorescence histograms for two of the strains are shown in Fig. 1b and the complete data are summarized in Table 4. Some of the most

striking differences were seen when H-2^k mice were examined (first panel of Table 4). All the strains in this set expressed a high frequency of V_{β8.2} T cells. Strains carrying the non-stimulatory Mls^b allele also contained a high frequency of V_{β8.1} T cells; however, strains expressing Mls^a lacked T cells expressing V_{β8.1} almost completely. This linkage of V_{β8.1} expression to Mls was particularly apparent when comparing CBA/J to CBA/Ca. These strains are nearly identical, yet CBA/Ca (Mls^b, H-2^k) had a high frequency of V_{β8.1} T cells and CBA/J (Mls^a, H-2^k) virtually no V_{β8.1} T cells.

These results correlate well with our results with V_{β8.1} T-cell hybridomas and indicate that the high frequency with which V_{β8.1} T cells recognize Mls^a results in the elimination of these cells by the mechanisms of self-tolerance. The depletion of V_{β8.1} T cells in Mls^a mice was perhaps even greater than expected, as 20% of V_{β8.1} T cell hybridomas from B10.BR (Mls^b, H-2^k) mice could not be shown to be Mls^a-reactive. These results indicate that clonal elimination during induction of self-tolerance may be a more sensitive means of detecting the reactivity of an αβ receptor than is the stimulation of a T-cell hybridoma.

The influence of Mls^a on the level of V_{β8.1} T cells was confirmed when non-H-2^k mice were examined. A series of H-2 congenic mice on the B10 background (Mls^b) were compared to the H-2 congenic mice on the D2 or D1 background (Mls^a) described above. The results are summarized in the second panel of Table 4. All the mice had high levels of V_{β8.2} T cells. As was seen with H-2^k mice, the Mls^b allele allowed a normal frequency of V_{β8.1} T cells. Mls^a again led to suppression of V_{β8.1} T cells; however, there was a hierarchy of expression depending on H-2 type. DBA/2 (H-2^d) mice were nearly as suppressed as H-2^k mice. D1.LP (H-2^b) mice were intermediate and DBA/1 (H-2^a) and D2.GD (recombinant H-2^d/H-2^b) mice appeared normal in their level of V_{β8.1} T cells. These results complemented our findings with V_{β8.1} T cell hybridomas, and showed that, when specificity was measured by clonal elimination during tolerance induction, only some H-2 products effectively combined with Mls^a to create a ligand which could be recognized by V_{β8.1} T cells.

We examined two other types of mice to strengthen our conclusion that the low level of V_{β8.1} T cells in Mls^a mice was due to self-tolerance, rather than to differences among the V_{β8.1} genes of these strains. First, we examined F₁ mice between C3H/HeJ (Mls^b, H-2^k) and DBA/2 (Mls^a, H-2^d). If the lack of V_{β8.1} expression in DBA/2 mice was due to tolerance to Mls^a/H-2^d, then the high expression of V_{β8.1} in C3H/HeJ should have been suppressed in the F₁ mice. If the low expression were intrinsic to the DBA/2 V_{β8.1} gene, then the expression in the F₁ should have been intermediate between that in C3H/HeJ and DBA/2. As seen in panel three of Table 4, the former prediction was correct. The level of expression of V_{β8.1} in the F₁ mice was

Table 2 V_{β8.1} T-cell hybrids from B10.BR (Mls^b, H-2^k) mice react to Mls^a with high frequency

V _{β8} use by KJ16 ⁺ hybrids	No. of hybrids	Hybrids reactive to AKR/J and CBA/J (H-2 ^k , Mls ^a)		
		No. tested	No. reactive	Per cent reactive
1	27	19	15	79
2	50	30	4	13

KJ16⁺ T-cell blasts were prepared from B10.BR (Mls^b, H-2^k) lymph-node T cells as previously described²⁵. T-cell hybrids were prepared by fusing^{39,40} these blasts to a variant of the T-cell thymoma BW5147, which had radiation-induced deletions of its functional V_α or V_β genes (White *et al*, manuscript in preparation). Hybrids were originally screened with KJ16 and F23.2 for V_{β8} usage. A portion of those determined to be either V_{β8.1} or V_{β8.2} were tested for IL-2 production^{25,39,40} in response to CBA/J or AKR/J (Mls^a, H-2^k) spleen cells. The limit of detection was 10 units of IL-2 ml⁻¹.

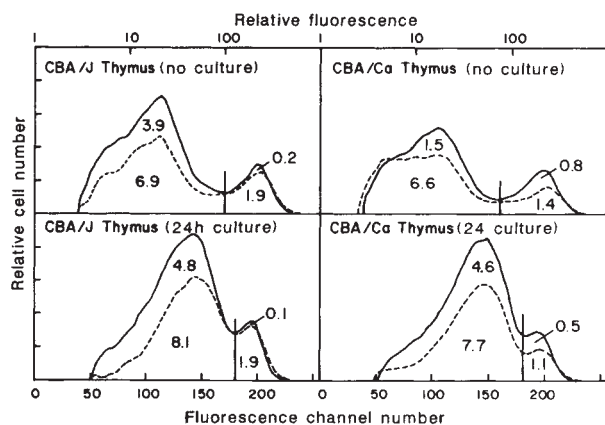


Fig. 2 Mls eliminates $V_{\beta 8.1}^+$ mature, but not immature, thymocytes. Thymocytes were prepared from either CBA/Ca(Mls^b, H-2^k) or CBA/J(Mls^a, H-2^k) mice. The cells were stained with KJ16 and F23.2 either immediately after preparation or after 24 h in suspension culture and analysed as in Fig. 1. Similarly prepared and stained thymocytes from C57BR mice, which lack the structural genes for the $V_{\beta 8}$ family, were used in each case as a negative control. Two hundred thousand cells were analysed in each sample. The histograms shown were derived by computer-aided subtraction of the appropriate normalized negative control histogram from the experimental histograms. For comparison the results with KJ16 (solid line) and F23.2 (dashed line) are overlaid on the same histogram. Also shown are the calculated percentage of the thymocytes expressing either $V_{\beta 8.1}$ (upper numbers) or $V_{\beta 8.2}$ (lower numbers) within the dully-staining, immature population (left region) or brightly-staining, mature population (right region)^{7,8}.

as low as that in DBA/2, indicating the dominant suppression by Mls^a predicted by tolerance. Finally, we examined a set of recombinant inbred mice³⁰ between AKR (Mls^a, H-2^k) and C57L (Mls^b, H-2^b). C57L lack the structural genes for the $V_{\beta 8}$ family²² and, therefore, only those mice in this set whose V_{β} complex came from AKR are capable of $V_{\beta 8}$ expression. Previously²⁷, we have shown that KJ16 expression in these mice is lower in the presence of Mls^a than Mls^b. We re-examined six of these strains to determine $V_{\beta 8.1}$ and $V_{\beta 8.2}$ expression (Table 4, fourth panel). All six strains expressed high levels of $V_{\beta 8.2}^+$ T cells. The four Mls^b strains had high levels of $V_{\beta 8.1}^+$ T cells, the Mls^a/H-2^k strain had virtually no $V_{\beta 8.1}^+$ T cells and, as before, the Mls^a/H-2^b strain had intermediate levels of $V_{\beta 8.1}^+$ T cells. Thus the AKR $V_{\beta 8.1}$ gene is functional when bred into a mouse lacking the suppressive Mls^a allele, supporting the conclusion that the low levels of $V_{\beta 8.1}^+$ T cells are due to clonal deletion during tolerance induction.

During the induction of self-tolerance to the class II H-2 molecule, I-E, most $V_{\beta 17a}^+$ T cells are deleted at the transition between immature and mature cells in the thymus. We examined whether tolerance to the combination of the antigen Mls^a plus H-2 products occurred to a similar stage in T-cell development. Thymocytes from young adult CBA/Ca(Mls^b, H-2^k) or CBA/J(Mls^a, H-2^k) mice were stained with KJ16 and F23.2. We have previously shown that at the time cells are removed from the thymus the density of $\alpha\beta$ receptor on immature cortical thymocytes is ~10-fold lower than on mature medullary thymocytes or peripheral T cells^{7,8,26,31}. After 24 h culture in suspension, receptor density on immature thymocytes spontaneously increases about three fold, whereas receptor density on mature cells does not change significantly^{7,8}. Therefore, we distinguished mature cells from immature cells by the relative intensity of fluorescence, staining thymocytes at the time of preparation and again after 24 h in suspension culture. Fluorescence histograms and the calculated proportion of $V_{\beta 8.1}^+$ and $V_{\beta 8.2}^+$ mature and immature thymocytes are shown in Fig. 2. The results show that $V_{\beta 8.1}$ is expressed well in both strains on

Table 3 Mls^a-reactive KJ16⁺ T-cell hybrids are H-2 restricted

Strain	IL-2 production (units ml ⁻¹) in response to							
	B10.BR	C3H	CBA/J	AKR	DBA/2	D1.LP	D2.GD	DBA/1
	H-2 Mls	kkkk b	kkkk b	kkkk a	kkkk a	dddd a	bb—b a	dd—b a
KJ16 ⁺ hybrid								
V ⁺ _{β8.1}								
K16-20	—	—	>810	>810	>810	>810	—	—
K16-70	—	—	270	810	270	270	—	—
K16-57	—	—	810	810	270	270	—	—
K16-33	—	—	810	810	270	270	—	—
K16-12	—	—	470	810	90	90	—	—
K16-46	—	—	10	10	10	10	—	—
K16-42	—	—	810	810	30	270	—	—
K16-55	—	—	>810	>810	10	90	—	—
K16-63	—	—	810	810	810	10	—	—
K16-67	—	—	270	810	270	30	—	—
K16-66	—	—	>810	>810	90	30	—	—
K16-35	—	—	90	90	90	10	—	—
V ⁺ _{β8.2}								
K16-15	—	—	>810	>810	—	—	—	—
K16-41	10	10	90	90	10	—	—	—
K16-36	—	—	50	50	—	—	—	—
K16-45	—	—	30	30	—	—	—	—

The Mls-reactive KJ16⁺ hybrids described in Table 2 were tested for IL-2 production in response to spleen cells from different strains of mice. For each strain the Mls allele and the expressed allelic form of the H-2K, H-2A, H-2E and H-2D molecules respectively are shown. Twelve of the 15 $V_{\beta 8.1}$ hybrids and the four $V_{\beta 8.2}$ hybrids were analysed.

immature thymocytes, but expression on mature thymocytes is much lower in CBA/J than in CBA/Ca. Thus, tolerance to the self antigen Mls^a plus products of H-2 occurs within the thymus at the transition between the immature cortex and the mature medulla.

Discussion

Our results have implications for understanding T cell/antigen interactions in general and the phenomenology associated with the Mls locus in particular. Mls has been a thorny problem for immunologists for many years. The durable interest in Mls (despite its demonstration only in the mouse) stems from the conviction that understanding how it stimulates such a high proportion of T cells will reveal something fundamental about T-cell antigen recognition. A major controversy has been whether Mls^a stimulates T cells via the $\alpha\beta$ receptor or through a separate, less variable T-cell receptor. The extremely high frequency of responding T cells¹³⁻¹⁶, the existence of T-cell clones that respond to Mls^a on cells of foreign H-2 type^{19,32,33} and the apparent segregation of Mls^a reactivity and foreign antigen/MHC recognition in some T-cell hybridomas carrying both reactivities²⁰ have all been presented as suggesting a non- $\alpha\beta$ receptor for Mls^a. On the other hand, the existence of T-cell clones that discriminate the H-2 type of Mls^a-presenting cells^{17,18} and the inhibition of Mls^a reactivity by antibodies to class II MHC products and by antibodies to the T-cell $\alpha\beta$ receptor³³ have provided evidence for involvement of MHC and the $\alpha\beta$ receptor in Mls^a recognition. Obviously, therefore, the usefulness of Mls^a as a model antigen depends on the resolution of this controversy. Our data establish the linkage of V_{β} usage and H-2 type to Mls^a recognition, strongly favouring the conclusion that Mls^a is recognized by the $\alpha\beta$ receptor as an antigen associated with MHC.

The ability of the Mls^a/MHC ligand to be recognized by such a large proportion of T cells, including virtually all $V_{\beta 8.1}^+$ T cells, provides a tool to follow this population through T-cell development in the presence or absence of the ligand. For example, there has been considerable controversy over the mechanism whereby tolerance is established to self-antigens. Although the deletion of self-reactive T-cell clones appears as the simplest hypothesis, some experiments have suggested that

mechanisms exist that control the activation of self-reactive T cells without their elimination³⁴. The virtual disappearance of $V_{\beta 8.1}^+$ cells during T-cell maturation in Mls^a mice of appropriate H-2 type is a clear demonstration of the deletion of self-reactive clones as one means of self-tolerance. This conclusion was also supported by our previous results^{25,26} using $V_{\beta 17a}^+$ T cells, most of which recognize the class II molecule, I-E, and are eliminated at a similar stage of T-cell development in I-E-bearing mice.

How the T-cell $\alpha\beta$ receptor interacts with an antigen/MHC ligand is still unknown. It is reasonable to predict that, given the similarity among the ligands for this receptor, there will prove to be a general geometry to the tri-molecular complex with multiple contact points between the variable elements of the receptor and the antigen and MHC in the ligand contributing to its overall stability. The recently-reported three-dimensional structure of the human class I MHC molecule, HLA-A2, is consistent with this view^{5,6}. Unfortunately, it has been difficult to analyse this problem by examining the receptor-variable elements used in the response to different defined antigen/MHC ligands. Even well-defined antigens usually invoke heterogeneous responses. Small changes in either the antigen or MHC component of the ligand can have large effects on the repertoire of responding T cells^{35,36}. In a few cases, where antigens have been defined which induce T-cell responses of limited heterogeneity, potential contact sites between receptor and ligand have been identified^{35,37,38}, but as yet there is not enough information to construct general rules.

Our studies with $V_{\beta 8.1}$ provide a different approach to the problem. In this case the analysis has been simplified because interactions between the ligand and the V_{β} portions of the receptor can be so strong as to be apparent almost regardless of the contributions of the other receptor components. For example, the presence of $V_{\beta 8.1}$ in a T-cell receptor virtually assures reactivity with Mls^a presented by H-2^k or H-2^d cells, indicating a strong interaction between this variable element of the receptor and the antigen portion of the ligand. But an interaction of $V_{\beta 8.1}$ with MHC as well is suggested by the differences in H-2 specificity between Mls^a-reactive T cells using either $V_{\beta 8.1}$ or $V_{\beta 8.2}$ (Table 3). An interaction between V_{β} and MHC was also suggested in our previous studies by the strong correlation between $V_{\beta 17a}$ usage and recognition of I-E-containing ligands^{25,26}.

Those cases where $V_{\beta 8.1}^+$ T cells failed to react to Mls^a were also informative. For example, we and others¹⁸ have failed to find T cells reactive to Mls^a presented by H-2^q T cells. This result could be most easily explained by the failure of Mls^a to associate effectively with products of H-2^q haplotypes much in the same way that many conventional antigens bind effectively only to particular H-2 products². In addition, although all $V_{\beta 8.1}^+$ T-cell hybrids from H-2^k mice responded to some degree to Mls^a on H-2^b cells, about one third of $V_{\beta 8.1}^+$ T cells survived tolerance induction in Mls^a/H-2^b mice (Table 3). Thus the repertoires of $V_{\beta 8.1}^+$ T cells in H-2^b and H-2^k mice are not identical (presumably reflecting the differences in H-2 products present during thymic selection), and imply that other variable components of the $V_{\beta 8.1}$ -bearing receptor (for example, J_{β} , V_{α} , J_{α}) influence which specific products of H-2 when combined with Mls^a can be recognized by that particular receptor. Overall, therefore, these results still indicate a rather complex interaction between the receptor and its ligand.

Finally the question can be raised whether degenerate ligands such as Mls^a/H-2 are simply anomalies providing useful tools to study $\alpha\beta$ receptor/ligand interactions or whether these types of ligands are important in T-cell function as well. Clearly, Mls^a-like molecules cannot be prevalent among self or foreign antigens, as then few T cells would survive self-tolerance and the specificity of the response of T cells would be lost. But there may be a role for these types of ligands in the positive selection of the T-cell repertoire during interaction with self-MHC on the thymic epithelium. The nature of the self-antigens associated

Table 4 Control of $V_{\beta 8.1}$ expression by Mls^a and H-2

Strain	H-2				Mls	% of peripheral T cells expressing	
	K	A	E	D		$V_{\beta 8.1}$	$V_{\beta 8.2}$
C3H/HeJ	k	k	k	k	b	8.2±0.2	15.9±0.2
B10.BR	k	k	k	k	b	6.5±0.4	12.4±0.4
CBA/Ca	k	k	k	k	b	6.2±0.1	12.5±0.1
CBA/J	k	k	k	k	a	0.2±0.1	14.5±0.1
AKR/J	k	k	k	k	a	0.3±0.1	11.9±0.2
RF/J	k	k	k	k	a	0.2±0.1	12.0±0.1
B10.D2	d	d	d	d	b	6.4±0.1	13.2±0.0
C57BL/10	b	b	-	b	b	6.2±0.2	10.4±0.1
B10.Q	q	q	-	q	b	4.3±0.2	11.4±0.3
DBA/2	d	d	d	d	a	0.6±0.0	14.9±0.1
D1.LP	b	b	-	b	a	1.7±0.1	12.2±0.3
D2.GD	d	d	-	b	a	4.0±0.1	14.9±0.2
DBA/1	q	q	-	q	a	4.9±0.2	11.9±0.1
C3D2F ₁	k	k	k	k	b	0.5±0.2	17.6±0.4
	d	d	d	d	a		
AKXL-19	b	b	-	b	b	5.7, 5.7	10.2, 9.9
AKXL-13	k	k	k	k	b	6.0, 6.0	12.1, 11.8
AKXL-14	k	k	k	k	b	7.6, 7.4	12.0, 11.7
AKXL-21	k	k	k	k	b	8.2, 8.4	11.7, 11.8
AKXL-17	b	b	-	b	a	2.7, 2.7	11.1, 10.9
AKXL-28	k	k	k	k	a	0.3, 0.3	11.6, 11.5

Purified lymph node T cells from the mouse strains indicated were analysed for binding of KJ16 and F23.2 as in Fig. 1. These results were used to calculate the percent of $V_{\beta 8.1}^+$ and $V_{\beta 8.2}^+$ T cells. The average percentages±s.e.m. are shown based on analysis of three separate mice on different days, except for the AKXL recombinant inbred mice, where only two mice from each strain were analysed and both determinations are shown. All mice were either purchased from the Jackson Laboratories, Bar Harbor, or bred in our own facilities.

with MHC during this process is unknown. If they are identical to those associated with MHC on the cells which induce self-tolerance, one is left with the problem of how any positively selected T cells escape tolerance induction. Although differences in the affinity of interaction required for positive selection as opposed to tolerance may solve this problem, another possibility is that the MHC ligands involved in positive selection are not identical to those involved in tolerance induction. For example, degenerate ligands with properties similar to Mls^a/MHC may be uniquely expressed on the thymus epithelium. By interacting ubiquitously with relatively invariant residues in that part of the $\alpha\beta$ receptor which usually contacts the antigen portion of an antigen/MHC complex, these degenerate ligands could promote selection on the basis of weak interaction sites on the rest of the receptor with MHC residues, accounting both for the skewing of the repertoire towards antigen/self-MHC ligands and for the survival of selected cells through the process of self-tolerance.

We thank Jerome Bill and David Diguisto for providing the BBM series of hybridomas in Table 1 and helping in their analysis, Edward Palmer for providing the probes used in Southern blot analyses, and Ella Kushnir and Terri Wade for excellent technical assistance. Supported in part by research grants from the NIH and the American Cancer Society.

Received 5 October; accepted 31 December 1987.

- Babbitt, B., Allen, P., Matsueda, G., Haber, E. & Unanue, E. *Nature* **317**, 359-361 (1985).
- Buus, S., Sette, A., Colon, S., Jenis, D. & Grey, H. *Cell* **47**, 1071-1077 (1986).
- Zinkernagel, R. & Doherty, P. *Nature* **251**, 547-548 (1974).
- Bevan, M. *et al. Immun. Reviews* **42**, 1-270 (1978).
- Bjorkman, P. *et al. Nature* **329**, 506-512 (1987).
- Bjorkman, P. *et al. Nature* **329**, 512-518 (1987).
- Born, W. *et al. J. Immun.* **138**, 999-1008 (1987).
- McDuffie, M., Born, W., Marrack, P. & Kappler, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8728-8733 (1986).

9. Kruisbeek, A. *et al.* *J. exp. Med.* **161**, 1029-1047 (1985).
10. Bevan, M. *Nature* **269**, 417-419 (1977).
11. Zinkernagel, R. *et al.* *J. exp. Med.* **147**, 882-896 (1978).
12. Billingham, R., Brent, L. & Medawar, P. *Nature* **172**, 603-606 (1956).
13. Festenstein, H. *Transplant. Rev.* **15**, 62-88 (1973).
14. Festenstein, H. *Transplantation* **18**, 555-557 (1974).
15. Festenstein, H., Bishop, C. & Taylor, B. *Immunogenetics* **5**, 357-361 (1977).
16. Janeway, C., Lerner, E., Jason, J. & B. Jones *Immunogenetics* **10**, 481-497 (1980).
17. Jones, B. & Janeway, C. *Immunogenetics* **16**, 243-256 (1982).
18. Lynch, D., Gress, R., Needleman, B., Rosenberg, S. & Hodes, R. *J. Immun.* **134**, 2071-2078 (1985).
19. Webb, S., Molnar-Kimber, K., Jennifer, B., Sprent, J. & Wilson, D. *J. exp. Med.* **154**, 1970-1974 (1981).
20. Webb, S. *et al.* *J. exp. Med.* **161**, 269-274 (1985).
21. Staerz, U., Rammensee, H., Benedetto, J. & Bevan, M. *J. Immun.* **134**, 3994-4000 (1985).
22. Behlke, M. *et al.* *J. exp. Med.* **165**, 257-262 (1987).
23. Haskins, K. *et al.* *J. exp. Med.* **160**, 452-471 (1984).
24. Staerz, U. & Bevan, M. in *Molecular Biology of the Immune System* (eds Streilein, J., Ahmad, F., Black, S., Blomberg, B. & Voellmy, R.) 61-64 (Cambridge University Press, Cambridge, 1985).
25. Kappler, J. *et al.* *Cell* **49**, 263-271 (1987).
26. Kappler, J., Roehm, N. & Marrack, P. *Cell* **49**, 273-280 (1987).
27. Roehm, N. *et al.* *J. Immun.* **135**, 2176-2182 (1985).
28. Abe, R., Ryan, J. & R. Hodes, *J. exp. Med.* **165**, 1113-1129 (1987).
29. Click, R. *et al.* *J. Immun.* **139**, 321-325 (1987).
30. Taylor, B. in *Genetic Variants and Strains of the Laboratory Mouse* (ed. Green, M.) 397 (Fisher, Stuttgart, 1981).
31. Roehm, N. *et al.* *Cell* **38**, 577-584 (1984).
32. Glasebrook, A. & Fitch, F. *J. exp. Med.* **151**, 876-895 (1980).
33. Katz, M. & Janeway, C. *J. Immun.* **134**, 2064-2070 (1985).
34. Hasek, M. *et al.* *Immun. Rev.* **46**, 1-86, 123-146 (1979).
35. Allen, P., McKean, D., Beck, B., Sheffield, J. & Glimcher, L. *J. exp. Med.* **162**, 1264-1274 (1985).
36. Ogasawara, K., Maloy, W. & Schwartz, R. *Nature* **325**, 450-452 (1987).
37. Fink, P., Matis, L., McElligott, D., Bookman, M. & Hedrick, S. *Nature* **321**, 219-226 (1986).
38. Winoto, A. *et al.* *Nature* **324**, 679-682 (1986).
39. Kappler, J., Skidmore, B., White, J. & Marrack, P. *J. exp. Med.* **153**, 1198-1214 (1981).
40. Marrack, P. in *Isolation, Characterization and Utilization of T-Lymphocyte Clones* (eds Fathman, G. & Fitch, F.) 508-510 (Academic, New York, 1982).
41. Kranz, D., Sherman, D., Sitkovsky, M., Pasternack, M. & Eisen, H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 573-577 (1984).
42. Kaye, J., Porcelli, S., Tite, J., Jones, B. & Janeway, C. *J. exp. Med.* **158**, 836-856 (1983).
43. White, J., Haskins, K., Marrack, P. & Kappler, J. (1983) *J. Immun.* **130**, 1033-1037 (1983).
44. Blackman, M. *et al.* *Cell* **47**, 349-357 (1986).
45. Endres, R., Marrack, P. & Kappler, J. (1983) *J. Immun.* **131**, 1656-1662 (1983).
46. Choi, H. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 1992-1996 (1987).

T-cell receptor V_{β} use predicts reactivity and tolerance to Mls^a -encoded antigens

H. Robson MacDonald*, Reto Schneider†, Rosemary K. Lees*, Rawleigh C. Howe*, Hans Acha-Orbea‡, Hilliard Festenstein‡, Rolf M. Zinkernagel† & Hans Hengartner†

* Ludwig Institute for Cancer Research, Lausanne Branch, 1066 Epalinges, Switzerland

† Institute of Pathology, University Hospital, 8091 Zurich, Switzerland

‡ Department of Immunology, London Hospital Medical College, London, UK

T lymphocytes reactive with the product of the Mls^a -allele of the minor lymphocyte stimulating (Mls) locus use a predominant T-cell receptor β -chain variable gene segment ($V_{\beta 6}$). Such $V_{\beta 6}$ -bearing T cells are selectively eliminated in the thymus of Mls^a -bearing mice, consistent with a model in which tolerance to self antigens is achieved by clonal deletion.

THE mechanism(s) underlying the selection of the T-cell repertoire of antigenic specificities remains one of the most fundamental unanswered questions in immunology. It is now clear that the genes encoding the α and β chains of the major histocompatibility complex (MHC)-restricted heterodimeric T-cell antigen receptor (TCR) are rearranged and expressed in the microenvironment of the developing thymus¹. Furthermore, it is generally accepted that the observable (that is, post-thymic) repertoire is shaped by negative selection events² (to explain self:nonself discrimination or 'tolerance') as well as positive selection³ (to explain MHC-restricted antigen recognition). But the cellular and molecular basis of repertoire selection remains obscure.

A major technical problem in the analysis of repertoire development has been the necessity of using complicated functional assays to evaluate the presence or absence of T cells of a given specificity. This limitation has however been overcome in a recent study by Kappler *et al.*^{4,5}, who identified a TCR β -chain variable gene segment (designated $V_{\beta 17a}$) that was preferentially used in panel of T-cell hybridomas reacting with the MHC class II gene product I-E^k. By exploiting a polymorphism in this locus that was recognized by a monoclonal antibody (mAb), it was possible to show that the development of receptors using $V_{\beta 17a}$ was suppressed in the thymus of mice expressing I-E^k. Thus a direct link between TCR specificity and elimination of self-MHC-reactive cells (putatively tolerance) could be established.

In contradistinction to the results of Kappler *et al.*⁴, many investigators have found either no correlation or conflicting data with regard to the use of particular TCR α - or β -chain variable domains in the T-cell response to various MHC-restricted antigens⁶⁻¹³. Thus it is currently not clear whether the $V_{\beta 17a}$ correlation with I-E^k represents an isolated observation or a useful paradigm.

In the course of screening peripheral T cells from a panel of inbred mouse strains with a new monoclonal antibody that recognizes all receptors that use the $V_{\beta 6}$ gene segment, we made the unexpected observation that $V_{\beta 6}$ usage correlates negatively with expression of the Mls^a allele of the minor lymphocyte-stimulating (Mls) locus¹⁴. Although the structure and function of the Mls gene product(s) is unknown, Mls^a is remarkable in that it is the only known antigen other than the MHC that can specifically stimulate proliferation of a high proportion of T lymphocytes from unimmunized mice¹⁵⁻¹⁸. Further analysis revealed that reactivity *in vitro* to Mls^a is strongly correlated with $V_{\beta 6}$ expression and that $V_{\beta 6}$ T cells are eliminated intrathymically in mice that express Mls^a . These data thus support the concept that physiological tolerance is mediated via the intrathymic elimination of self-reactive cells rather than by alternative mechanisms such as anergy or peripheral suppression. Furthermore, they have interesting implications for the nature of the Mls^a gene product and for the role of germline V_{β} segments in the selection of the T-cell repertoire.

Expression of $V_{\beta 6}$

We have recently shown that 44-22-1, a mAb raised originally against a H-2D^b-specific cytolytic T cell clone¹⁹, in fact recognizes all TCR using the $V_{\beta 6}$ gene segment²⁰. As shown in Fig. 1, 11% of lymph node cells from BALB/c mice stained with the 44-22-1 antibody, as compared to 18% with KJ16-133 (ref. 21; a mAb that recognizes TCR using $V_{\beta 8.1}$ or $V_{\beta 8.2}$, see ref. 22). By two-colour immunofluorescence, a similar proportion of CD4⁺ (14%) and CD8⁺ (12%) cells stained with 44-22-1, indicating that both subsets of mature T lymphocytes use the $V_{\beta 6}$ gene product to a similar extent. Control double-staining with KJ16 confirmed earlier studies²³ showing a slightly more frequent $V_{\beta 8}$ use in the CD8⁺ (30%) compared to the CD4⁺ (24%) subset. Interestingly, the intensity of 44-22-1 (or KJ16)

Table 1 Expression of $V_{\beta 8}$ and $V_{\beta 6}$ by T cells from inbred mouse strains

Strain	H-2	Mls	Lymph node		Cortisone-resistant thymus	
			KJ16 ($V_{\beta 8}$)	44-22-1 ($V_{\beta 6}$)	KJ16	44-22-1
C57Bl/6	b	b	12.4	5.8	17.7	8.5
BALB/c	d	b	18.0	11.0	24.9	11.7
DBA/2	d	a	12.1	1.0	20.0	1.5
BIO.G	q	b	8.9	4.1	ND	ND
A/J	k/d	c	11.4	9.1	ND	ND
CBA/J	k	d	10.9	0.4	23.6	1.3
CBA/Ca	k	b	20.8	14.2	ND	ND
C3H/He	k	c	13.0	10.5	ND	ND
C3H/Ou	k	c	14.9	13.8	25.3	13.2
C3H/Tif	k	a	10.7	0.6	ND	ND
AKR/J	k	a	6.8	0.6	ND	ND
BIO.BR	k	b	8.8	5.2	20.9	11.1
SJL	s	c	0.0	8.2	ND	ND
(C57Bl/6 × CBA/J) F_1	b × k	b × d	10.4	0.2	ND	ND
(C57Bl/6 × DBA/2) F_1	b × d	b × a	5.9	0.5	ND	ND
(C57Bl/6 × BALB/c) F_1	b × d	b × b	15.7	5.6	ND	ND
(DBA/2 × CBA/J) F_1	d × k	a × d	6.8	0.8	ND	ND
(DBA/2 × BALB/c) F_1	d × d	a × b	9.3	0.4	ND	ND
(BIO.G × BIO.BR) F_1	q × k	b × b	9.7	5.0	ND	ND
(BIO.G × DBA/2) F_1	q × d	b × a	10.3	0.7	ND	ND
(BIO.G × CBA/J) F_1	q × k	b × d	10.6	0.8	ND	ND

Lymph node cells or cortisone-resistant thymocytes (obtained 48 h after a single i.p. injection of 4 mg hydrocortisone acetate) were stained by indirect immunofluorescence with mAbs KJ16 (anti- $V_{\beta 8}$) or 44-22-1 (anti- $V_{\beta 6}$) and analysed by flow microfluorometry (as in Fig. 1). Data are presented as % positive cells (after subtraction of staining with the fluorescent conjugate alone). Mls typing was taken from the available literature⁴³ with the exception of C3H/Tif (H. Prah, personal communication). ND, not determined.

staining was slightly greater in the $CD4^+$ (compared to $CD8^+$) subset, a phenomenon initially noted by Roehm *et al.*²³. As light scatter analysis indicates that $CD4^+$ T cells are slightly smaller than $CD8^+$ T cells²⁴, the density of TCR expression (at least insofar as $V_{\beta 6}$ and $V_{\beta 8}$ are concerned) is greater for the $CD4^+$ subset of T lymphocytes.

Lack of TCR using $V_{\beta 6}$ in Mls^a mice

We next examined peripheral T cells from different mouse strains for their use of $V_{\beta 6}$. A summary of these data (Table 1) shows a surprising variability in the expression of 44-22-1. In contrast, expression of the $V_{\beta 8}$ gene family (as detected by mAb KJ16-133) was found in all strains except SJL (where genomic deletion of the $V_{\beta 8}$ family has occurred)²⁵. Inspection of the pattern of $V_{\beta 6}$ expression indicated no correlation with *MHC* gene products;

however, a striking negative correlation was seen with the expression of *Mls*. All strains tested that expressed either the Mls^a or Mls^d alleles had very few ($\leq 1\%$) 44-22-1⁺ cells, whereas strains expressing the Mls^b or Mls^c alleles had significant numbers (4–14%). Furthermore, analysis of a panel of F_1 hybrid mice (Table 1) indicated that absence of $V_{\beta 6}$ -bearing cells correlated with expression of Mls^a (or Mls^d) gene products by either one of the parental strains.

We therefore carried out further genetic mapping studies. In a first approach, two independent series of recombinant inbred (RI) strains ($Mls^b \times Mls^a$) were typed for expression of 44-22-1. All RI strains tested showed an inverse correlation between Mls^a expression and reactivity with mAb 44-22-1 (Table 2). These data were extended using a Mls^a congenic strain of BALB/c mice (BALB.D2. Mls^a) derived by 13 backcross gener-

Fig. 1 Expression of $V_{\beta 6}$ and $V_{\beta 8}$ by subsets of T lymphocytes of BALB/c mice. In *a* and *b* fresh lymph node cells were stained with rat mAbs 44-22-1 (anti- $V_{\beta 6}$) or KJ16-133 (anti- $V_{\beta 8}$) followed by fluorescent goat anti-rat immunoglobulin. In the lower panels, this staining was followed sequentially by biotinylated mAbs directed against mouse $CD4$ (*c*, *d*) or $CD8$ (*e*, *f*) and finally phycoerythrin-conjugated avidin. In all instances, samples were analysed on a FACS II flow cytometer equipped with a logarithmic fluorescence amplifier.

Methods. Aliquots of 10^6 cells in 100 μ l were stained at 4 °C as described in detail previously⁴⁷. Antibodies 44-22-1 (ref. 19) and KJ16 (ref. 21) were used as undiluted hybridoma culture supernatants. Rat mAbs GK-1.5 (anti- $CD4$) and 53-6.7 (anti- $CD8$) were biotinylated according to standard techniques⁴⁸ and used at 20 μ g ml^{-1} and 50 μ g ml^{-1} , respectively. Goat-anti-rat immunoglobulin (absorbed on mouse immunoglobulin) was provided by TAGO, Burlingame, CA. Avidin-phycoerythrin was prepared as previously described⁴⁹ and used at 100 μ g ml^{-1} . For flow cytometric analysis, viable cells were gated by a combination of narrow-angle forward light scatter and perpendicular light scatter⁵⁰. Histograms represent 10^4 viable cells whereas two-dimensional cytograms represent 4×10^4 viable cells. Control samples in *a* and *b* were stained with the fluorescent conjugate alone (thin lines).

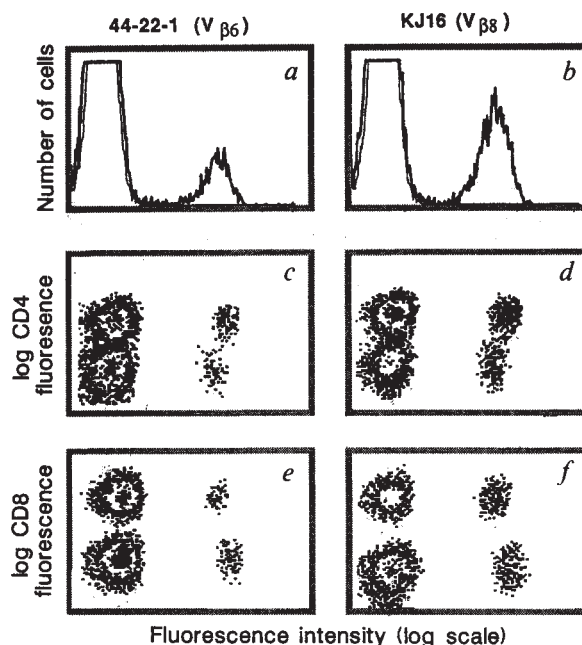


Table 2 Genetic mapping of lack of $V_{\beta 6}$ expression to Mls^a

Strain	H-2	Mls	KJ16($V_{\beta 8}$)	44-22-1 ($V_{\beta 6}$)
C57Bl/6	b	b	14.4	6.8
DBA/2	d	a	10.9	0.0
(B×D)F ₁	b×d	b×a	9.3	0.2
B×D	1	d	11.6	4.9
RI	2	b	6.0	1.2
	5	d	8.6	0.8
	6	d	16.6	8.9
	8	b	7.6	0.8
	9	d	7.2	1.3
	11	d	13.4	0.7
	12	d	13.1	6.7
	13	b	7.7	4.0
	14	b	6.6	6.9
	15	b	8.8	4.7
	16	d	12.5	7.1
	18	d	14.8	8.7
	19	b	7.3	3.4
	21	d	9.8	5.7
	22	d	10.6	0.3
	23	b	9.3	4.9
	24	d	8.5	0.2
	25	d	7.9	0.9
	27	d	7.3	0.5
	28	d	5.0	0.9
	29	b	9.5	0.7
	30	d	9.5	1.0
AK×L	6	k	0.6	9.9
RI	8	k	0.5	8.7
	13	k	11.1	6.2
	14	k	12.4	7.2
	16	b	0.3	5.5
	17	b	8.6	1.6
	19	b	9.4	5.7
	21	k	12.1	7.1
	24	b	0.2	8.2
	38	k	0.3	0.8
BALB/c	d	b	19.0±0.9*	10.0±1.0
BALB.D2.M1s ^a	d	a	13.6±0.7	0.4±0.4

Lymph node cells from C57Bl/6×DBA/2 (B×D) and AKR/J×C57L/J (AK×L) recombinant inbred (RI) strains or from Mls^a -congenic BALB/c mice²⁶ were stained and analysed as in Fig. 1. Data are presented as % positive cells (after subtraction of background staining with the fluorescent conjugate alone). Mls and H-2 typing of RI strains is taken from the literature^{27,38,44}. Negative KJ16 staining of some AK×L RI strains is due to $V_{\beta 8}$ deletion in parental C57L/J mice²⁵.

* Mean ± 1 s.d. of three individual mice.

ations²⁶. Table 2 shows that $V_{\beta 6}$ expression was ablated in the lymph nodes of the Mls^a congenic strain whereas $V_{\beta 8}$ expression was only moderately reduced. Taken together, these data establish formally that lack of $V_{\beta 6}$ expression maps with the Mls^a gene.

Intrathymic deletion of $V_{\beta 6}^+$ T cells

The thymus is the presumptive site of both T-cell development and tolerance induction. To ascertain whether mature T lymphocytes in the thymus of Mls^a mice also lack expression of $V_{\beta 6}$ determinants we treated a series of mice with hydrocortisone. The resulting cortisone-resistant thymocytes (CRT) consisted of only CD4⁺ or CD8⁺ T lymphocytes²⁴. As shown in Table 1, $V_{\beta 6}$ expression among CRT again correlated negatively with Mls^a , whereas $V_{\beta 8}$ expression was relatively constant. Thus the lack of expression of $V_{\beta 6}$ in the periphery of Mls^a strains also applies to thymocytes with a mature surface phenotype.

$V_{\beta 6}$ expression and Mls^a reactivity

The elimination of $V_{\beta 6}^+$ cells during intrathymic maturation in Mls^a mice could be seen as a manifestation of self-tolerance, assuming that this V_{β} segment confers preferential reactivity

towards the Mls^a product. This postulate was tested *in vitro* in several ways. First, homogeneous populations of Mls^a -reactive T cells were obtained by repeated stimulation of BALB/c (H-2^d, Mls^b) lymph node cells with irradiated DBA/2 (H-2^d, Mls^a) or congenic BALB.D2. Mls^a spleen cells. As no anti- Mls^a proliferative response of CD8⁺ T lymphocytes could be detected (even in the presence of excess IL-2; Fig. 2a), this and all subsequent experiments were carried out with CD4⁺ (that is, depleted of cells bearing CD8) responding cell populations. Table 3 shows that Mls^a -immune CD4⁺ cell populations were highly enriched for expression of $V_{\beta 6}$ (up to 79% after three stimulations). Parallel control (anti-MHC) cultures in which BALB/c T cells were re-stimulated with irradiated C57Bl/6 (H-2^b, Mls^b) spleen cells resulted in a much lower proportion (15–20%) of $V_{\beta 6}^+$ cells. Expression of $V_{\beta 8}$ determinants was similar in both anti- Mls^a and anti-MHC populations (15–20%). Thus $V_{\beta 6}$ -bearing cells are selectively enriched in Mls^a -immune populations generated *in vitro*.

In a second approach, $V_{\beta 6}^+$ cells were depleted from CD4⁺ BALB/c lymph node (by antibody plus complement treatment) and the resulting population was tested for reactivity to Mls^a or MHC antigens in a dose-response analysis. The results (Fig. 2b) demonstrated a threefold loss of anti- Mls^a reactivity in the $V_{\beta 6}^-$ population (compared to complement-treated control cells), whereas reactivity to MHC was unchanged. These data thus independently confirm that $V_{\beta 6}^+$ T cells account for ~2/3 of the Mls^a -responsive population in BALB/c mice.

Although these results firmly establish a correlation between $V_{\beta 6}$ expression and Mls^a reactivity, they did not directly show whether most (or all) $V_{\beta 6}^+$ T cells are potentially Mls^a -reactive. We therefore established a panel of T-cell hybridomas by fusing lectin-stimulated CD4⁺ BALB/c lymphocytes with the BW5147 thymoma and sorting the fused cells into CD4⁺ $V_{\beta 6}^+$ and CD4⁺ $V_{\beta 6}^-$ subsets. As shown in Fig. 3, most of the resulting cloned hybrids having functional TCR that use $V_{\beta 6}$ reacted specifically with Mls^a -bearing stimulator cells (11 of 14), whereas $V_{\beta 6}^-$ hybrids did not (0 of 22). Thus a high proportion of randomly-selected (lectin-stimulated) CD4⁺ T cells can recognize Mls^a determinants.

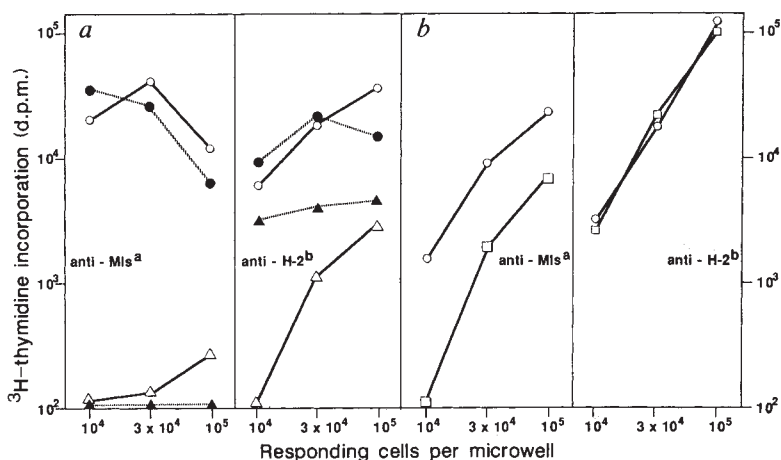
Discussion

Our data bear on two independent but clearly-related aspects of T-cell repertoire selection, namely the association of TCR V gene use with antigen and/or MHC specificity and the induction of tolerance to self antigens, and provide unequivocal evidence that expression of a particular V_{β} segment ($V_{\beta 6}$) confers preferential T-cell reactivity *in vitro* to Mls^a determinants and that TCR using $V_{\beta 6}$ are selected against in the thymus of Mls^a -expressing mice. Thus (at least in this model system) tolerance is mediated by the deletion of self-reactive clones rather than by alternative mechanisms such as anergy or peripheral suppression. Furthermore, our data provide the first clear-cut demonstration of an *in vivo* correlate of Mls^a reactivity.

The Mls locus¹⁴ is enigmatic in the sense that it is the only known genetic locus other than MHC in which a strong proliferative response is observed between unprimed individuals displaying different Mls 'alleles'. This locus is also peculiar in that the Mls response is not bidirectional: Mls^b mice respond to Mls^a , Mls^d and (to a much weaker extent) Mls^c , whereas mice bearing the other alleles do not respond to Mls^b . This has led to the suggestion that Mls^b is a silent allele. Furthermore, recent data from two laboratories^{27,28} indicate that the Mls^d allele can be considered to be a composite of Mls^a and Mls^c . Taken together with our data indicating that elimination of $V_{\beta 6}^+$ cells *in vivo* correlates with expression of Mls^a or Mls^d (but not Mls^c), the simplest view of the Mls system at present²⁹ would be that it is composed of a silent (Mls^b) allele and two independent stimulatory alleles (Mls^a and Mls^c). Although Mls^a has been mapped to chromosome 1 (ref. 30) and has been shown to be closely linked to a number of other genes^{31,32}, Mls^c has

Fig. 2 T-cell proliferation to *Mls^a*-encoded determinants. In *a*, nylon-wool purified BALB/c lymph node T cells were separated into CD4⁺ or CD8⁺ subsets by depletion with appropriate mAbs plus complement. ○, CD4⁺; ●, CD4⁺ + IL-2; △, CD8⁺; ▲, CD8⁺ + IL-2. In *b*, the CD4⁺ subset was further depleted of V_{β6} cells. ○, Total CD4⁺; □, V_{β6} CD4⁺. Various numbers of cells from each group were cultured with 5 × 10⁵ irradiated (2,000 rads) spleen stimulator cells differing from the responders at either *Mls^a* (BALB.D2.*Mls^a*) or H-2 (C57Bl/6) genetic loci. Recombinant human interleukin-2 (IL-2) was added at 60 ng ml⁻¹ where indicated. Proliferation was assayed by uptake of ³H-thymidine on day 4.

Methods. Complement-dependent depletion of CD4⁺ or CD8⁺ lymph node cells was accomplished by two consecutive treatments with rat IgM mAbs RL172.4 (anti-CD4) or 3.168.1 (anti-CD8), respectively⁵¹. Depletion of V_{β6} cells was carried out similarly using the rat IgM mAb 46-6B5 (ref. 19) that reacts with the same epitope (on TCR bearing V_{β6}) as 44-22-1 (ref. 20). Microcultures were established in flat-bottomed wells containing 200 μl Dulbecco's modified Eagle's medium supplemented with additional amino acids⁴⁵, 5% fetal bovine serum and 5 × 10⁻⁵ M 2-mercaptoethanol. ³H-thymidine (1 μCi per well) was added for 6 h on day 4 and samples were processed on an automated harvester (Cambridge Instruments). Data are presented as the mean d.p.m. of triplicate microcultures. Background proliferation (in the presence of irradiated syngeneic BALB/c spleen cells) has been subtracted. *a* and *b* are independent experiments.



not been mapped and could just as well be considered to be a completely independent gene (rather than an allele of *Mls*).

The biology of the *Mls^a* proliferative response is characteristic of class II MHC-restricted responses in that it is mediated exclusively by CD4⁺ T cells and can be blocked by either anti-CD4 or anti-class II MHC mAbs^{15,33-35}. The question of whether the *Mls^a* response can be considered to be MHC-restricted in the classical sense has been controversial^{15,36,37}, although studies using clonal T-cell populations have concluded that the *Mls^a* product is recognized in association with broadly cross-reactive (but still demonstrably polymorphic) class II MHC determinants³⁸. If correct, this interpretation leads to no simple explanation of the nature and role of the *Mls^a* gene product. Katz and Janeway³⁹ have postulated that *Mls^a*-reactive T cells bear TCR that recognize non-polymorphic class II MHC determinants with low affinity. According to their scheme, the *Mls^a* gene would code for some TCR-independent 'accessory' cell function that increases the TCR affinity for class II MHC to a level that allows effective triggering. Alternatively, it is possible that *Mls^a*-specific TCR recognize a processed product of the *Mls^a* gene presented by class II MHC molecules, as suggested by recent models of recognition of immunogenic peptides by T lymphocytes⁴⁰.

The degree to which TCR V genes can independently account for the recognition of MHC gene products and nominal antigens has been the subject of considerable debate⁶⁻¹³. Using panels of MHC-restricted T-cell clones and hybridomas, several groups have however observed correlations between V_β usage and MHC restriction elements^{8,13} and between V_α usage and nominal

antigens recognized⁹. More compellingly, the recent data of Kappler *et al.*⁴ indicate that one particular V_β segment (V_{β17a}) is frequently associated with TCR recognizing I-E^k. In our own studies, it is difficult to interpret the striking association of V_{β6} usage with *Mls^a* reactivity in the absence of any precise knowledge of the nature of the *Mls^a* gene product or its possible association (whether physical or functional) with MHC class II molecules. If however TCR of *Mls^a* reactive cells do have some intrinsic affinity for class II MHC molecules (as suggested by *in vitro* studies), then our data could be interpreted as providing a second dramatic example of preferential MHC recognition by a particular V_β gene product, strengthening the general case proposed for germline selection of MHC-reactive TCR V genes⁴¹.

Irrespective of its molecular explanation, the strong correlation observed between V_{β6} expression and *Mls^a* reactivity probably accounts for the high proportion of *Mls^a*-specific T lymphocytes in unprimed mice. Depending upon the genetic background 4-14% of peripheral lymph node cells (6-20% of total T lymphocytes) use the V_{β6} gene segment in *Mls^b* or *Mls^c* mice. Assuming that our panel of V_{β6}⁺ T cell hybrids is representative of the normal CD4⁺ T cell population, we would conclude that most V_{β6}⁺ CD4⁺ T cells respond to *Mls^a*, accounting for the 2-5% *Mls^a*-specific T cells defined by earlier limiting dilution studies^{18,42}.

Despite the overwhelming correlation between V_{β6} and *Mls^a*, it should nevertheless be pointed out that not all *Mls^a*-responsive T cells use this V segment. We consistently observed ~15% of V_{β8}⁺ T cells in polyclonal *Mls^a*-reactive populations generated

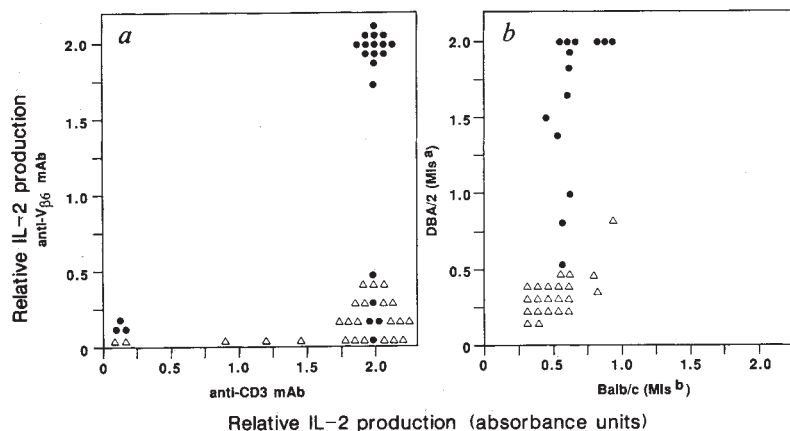
Table 3 Enrichment for V_{β6} cells in *Mls^a*-immune BALB/c populations

Responder	Mixed leukocyte culture Stimulator	Gene product recognized	% Positive cells	
			KJ16 (V _{β8})	44-22-1 (V _{β6})
BALB/c	B6 (1°)	H-2 ^b (+ minor H)	18.1	12.8
BALB/c	DBA/2 (1°)	<i>Mls^a</i> (+ minor H)	18.2	33.9
BALB/c	B6 (2°)	H-2 ^b (+ minor H)	23.6	15.9
BALB/c	DBA/2 (2°)	<i>Mls^a</i> (+ minor H)	19.1	48.4
BALB/c	B6 (3°)	H-2 ^b (+ minor H)	22.8	19.4
BALB/c	DBA/2 (3°)	<i>Mls^a</i> (+ minor H)	15.7	64.3
BALB/c	BALB.D2. <i>Mls^a</i> (1°)	<i>Mls^a</i>	19.9	61.8
BALB/c	BALB.D2. <i>Mls^a</i> (2°)	<i>Mls^a</i>	14.1	71.8
BALB/c	BALB.D2. <i>Mls^a</i> (3°)	<i>Mls^a</i>	11.0	78.7

Nylon-wool purified BALB/c lymph node cells were depleted of CD8⁺ cells by mAb + C' treatment (as in Fig. 2). Recovered cells (5 × 10⁵) were established in primary (1°) mixed leukocyte cultures⁴⁵ with 5 × 10⁶ irradiated (2,000 rads) C57Bl/6, DBA/2 or BALB.D2.*Mls^a* stimulator spleen cells for 5 days. For subsequent (2° and 3°) determinations long-term cultures were re-stimulated every 14 days and tested 4 days later⁴⁶. All cultures were analysed for expression of the determinants recognized by KJ16 and 44-22-1 as described in Fig. 1.

Fig. 3 Correlation of $V_{\beta 6}$ expression and Mls^a reactivity in CD4⁺ T-cell hybrids. A panel of T-T hybrids was generated by fusing CD8-depleted Concanavalin A (Con A)-stimulated BALB/c lymph node cells to BW5147 and sorting CD4⁺ hybrids on the basis of expression (or lack of it) of $V_{\beta 6}$ (Fig. 1c). After cloning by limiting dilution, 22 putative $V_{\beta 6}$ hybrids (●) and 24 negative hybrids (Δ) were screened for IL-2 production in the presence of immobilized anti- $V_{\beta 6}$ or anti-CD3 mAbs (a). Appropriate positive hybrids (14● and 22Δ) were then tested again for IL-2 production after stimulation with DBA/2 (Mls^a) or BALB/c (Mls^b) spleen cells (b).

Methods. BALB/c lymph node cells depleted of CD8⁺ cells (Fig. 2) were stimulated at 5×10^5 per ml with $2 \mu\text{g ml}^{-1}$ Con A for 3 days. Recovered cells were fused with a thioguanine-resistant variant of the BW5147 thymoma according to established procedures⁵² and selected in hypoxanthine, aminopterin, thymidine (HAT) medium for 7 days. The resulting population of putative hybrids was double-stained with mAbs 44-22-1 (anti- $V_{\beta 6}$) and GK-1.5 (anti-CD4) as in Fig. 1c. Sorted CD4⁺ $V_{\beta 6}$ and CD4⁺ $V_{\beta 6}$ populations were immediately cloned by limiting dilution (0.5 cells per well). To control for appropriate TCR expression of the cloned hybrids, purified mAbs 145-2C11 (anti-CD3)⁵³ or 46-6B5 (anti- $V_{\beta 6}$)¹⁹ were absorbed to round-bottomed microwells at a final concentration of $10 \mu\text{g ml}^{-1}$ and $2 \mu\text{g ml}^{-1}$, respectively. Cloned hybrid cells (10^5) were added in 200 μl to the plates in the presence of 1 ng ml^{-1} 12-O-tetradecanoyl phorbol-13 acetate (TPA). After 24 h, supernatants were harvested and assayed (at 1:10 final dilution) for IL-2 activity on IL-2-dependent CTLL-2 cells. Similarly, 10^5 cells from each hybrid clone were stimulated with 5×10^5 T-cell-depleted (anti-Thy-1.2 + C' treated) DBA/2 (Mls^a) or BALB/c (Mls^b) spleen cells and 24 h supernatants assayed for IL-2 at a 1:2 final dilution. All data are presented as absorbance units using an enzyme-linked immunosorbent assay (ELISA) modification of the CTLL assay⁵⁴. No significant IL-2 production was induced in any of the hybrids by TPA alone (data not shown).



in vitro. As cell numbers expanded over 100-fold in our anti-Mls^a cultures, non-Mls^a-responsive $V_{\beta 8}$ T cells should have been diluted out. Taken together with the *in vivo* data indicating a significant (although modest) reduction in the number of KJ16⁺ cells in congenic BALB.D2.Mls^a mice, it seems likely that one (or several) members of the $V_{\beta 8}$ gene family is also correlated with Mls^a responsiveness. At present, we cannot identify this putative correlation with precision; however the fact that mAb F23.1 (which reacts with all three members of the $V_{\beta 8}$ gene family)²² detects the same proportion of cultured Mls^a-reactive T cells as KJ16 (data not shown) would be consistent with the possibility that TCR using the $V_{\beta 8.1}$ and/or $V_{\beta 8.2}$ gene segments are also involved in the response to Mls^a.

A significant finding in the present study was the roughly equivalent expression of $V_{\beta 6}$ by both CD4⁺ and CD8⁺ subsets of mature T lymphocytes in Mls^b mice. These data are consistent with earlier studies of functional T-cell clones in which $V_{\beta 6}$ was shown to be used by both class I- and class II-MHC restricted cell lines⁶. In view of the fact that a very high proportion of randomly generated $V_{\beta 6}$ CD4⁺ T cell hybridomas are Mls^a reactive, one must find an explanation for the fact that Mls^a reactivity is not found among CD8⁺ T cells expressing $V_{\beta 6}$. One possibility would be that class I-restricted CD8⁺ $V_{\beta 6}$ T cells are selected for use of a distinct subset of TCR α -chain genes that are not compatible with Mls^a reactivity. Although it is cumbersome to test definitively, this hypothesis is not supported by the fact that the same V_{α} segment is used in functional $V_{\beta 6}$ helper (CD4⁺) and cytolytic (CD8⁺) T cell clones in at least one instance¹⁰. Alternatively, the accessory function of CD4 may be an obligate requirement for achieving detectable affinity for Mls^a determinants in cells that express $V_{\beta 6}$. This possibility is consistent with the observed high degree of susceptibility of Mls^a responses *in vitro* to inhibition by anti-CD4 antibodies³⁵. Functional experiments employing CD4⁻ loss variants of $V_{\beta 6}$ Mls^a-specific T cells or CD4⁺ transfectants of CD8⁺ $V_{\beta 6}$ T cells are required to test this latter hypothesis directly.

Despite their lack of demonstrable reactivity to Mls^a determinants *in vitro*, CD8⁺ $V_{\beta 6}$ T cells are nevertheless completely eliminated in the thymus of Mls^a-bearing mice. In developmental terms, such a finding implies strongly that tolerance induction must occur before (or independently of) the divergence of the CD4⁺ and CD8⁺ T-cell lineages. An attractive hypothesis to explain the elimination of $V_{\beta 6}$ CD8⁺ T cells would be to speculate that CD4⁺ CD8⁺ (cortical) thymocytes contain the precur-

sors of both CD4⁺ and CD8⁺ T-cell lineages and that tolerance induction (and perhaps also lineage commitment) is imposed at this stage of development. According to such a model, all $V_{\beta 6}$ cells would be eliminated in Mls^a mice at this stage because they simultaneously express CD4 and $V_{\beta 6}$ (see above). In contrast, in Mls^b or Mls^c mice, no such elimination would occur and further (putatively positive) selection mechanisms would account for the observed repertoire of $V_{\beta 6}$ cells among both CD4⁺ and CD8⁺ subsets. Alternative hypotheses in which $V_{\beta 6}$ cells would be tolerated in Mls^a mice at other stages of development (in the absence of CD4 and/or CD8) cannot of course be excluded; however they would only be compatible with the experimental data if one were to invoke a higher avidity of interaction between $V_{\beta 6}$ receptor-bearing cells and Mls^a determinants during the process of tolerance induction than is observed in the mature T-cell repertoire.

We thank P. Zaech and E. Santoli (Epalinges), M. Condrau, E. Niederer and M. Hirstand (Zurich) for performing the flow cytometry and P. Brunet for preparing the manuscript. Supported in part by a grant from the Swiss National Science Foundation (to H.H.).

Received 4 December 1987; accepted 7 January 1988.

1. Kronenberg, M., Siu, G., Hood, L. E. & Shastri, N. A. *Rev. Immun.* **4**, 529-591 (1986).
2. Nossal, G. J. V. A. *Rev. Immun.* **1**, 33-62 (1983).
3. Zinkernagel, R. & Doherty, P. C. *Adv. Immun.* **27**, 51-77 (1979).
4. Kappler, J. W. *et al. Cell* **49**, 263-271 (1987).
5. Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49**, 273-280 (1987).
6. Rupp, F., Acha-Orbea, H., Hengartner, H., Zinkernagel, R. & Joho, R. *Nature* **315**, 425-427 (1985).
7. Garman, R., Ko, J.-L., Vulpe, C. & Raulet, D. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3987-3991 (1986).
8. Fink, P. J., Matis, L. A., McElligott, D. L., Bookman, M. & Hedrick, S. M. *Nature* **321**, 219-226 (1986).
9. Winoto, A. *et al. Nature* **324**, 679-682 (1986).
10. Rupp, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 219-222 (1987).
11. Gorman, J. *et al. Cell* **40**, 859-867 (1985).
12. Iwamoto, A. *et al. J. exp. Med.* **165**, 591-600 (1987).
13. Morel, P., Livingstone, A. M. & Fathman, C. G. *J. exp. Med.* **166**, 583-588 (1987).
14. Festenstein, H. *Transplant. Rev.* **15**, 62-88 (1973).
15. Janeway, C. A., Lerner, E. A., Jason, J. M. & Jones, B. *Immunogenetics* **10**, 481-497 (1980).
16. Lutz, C. A., Glasebrook, A. L. & Fitch, F. W. *Eur. J. Immun.* **11**, 726-734 (1981).
17. Kronen, M., Scheurich, P., Pfizenmaier, K., Röllinghoff, M. & Wagner, H. *J. exp. Med.* **156**, 41-54 (1982).
18. Miller, R. A. & Stutman, O. *J. Immun.* **128**, 2258-2264 (1982).
19. Acha-Orbea, H., Zinkernagel, R. M. & Hengartner, H. *Eur. J. Immun.* **15**, 31-36 (1985).
20. Hengartner, H. *et al.* (in preparation).
21. Haskins, K. *et al. J. exp. Med.* **160**, 452-471 (1984).
22. Behlke, M. A. *et al. J. exp. Med.* **165**, 257-262 (1987).
23. Roehm, N. *et al. Cell* **38**, 577-584 (1984).
24. Ceredig, R., Djalynas, D. P., Fitch, F. W. & MacDonald, H. R. *J. exp. Med.* **158**, 1654-1671 (1983).

25. Sim, G. K. & Augustin, A. A. *Cell* **42**, 89-92 (1985).
26. Festenstein, H. & Berumen, L. *Transplantation* **37**, 322-324 (1984).
27. Abe, R., Ryan, J. J. & Hodes, R. J. *J. exp. Med.* **165**, 1113-1129 (1987).
28. Click, R. E. *et al. J. Immun.* **139**, 321-325 (1987).
29. Abe, R., Ryan, J. J. & Hodes, R. J. *J. exp. Med.* **166**, 1150-1155 (1987).
30. Festenstein, H., Bishop, C. & Taylor, B. A. *Immunogenetics* **5**, 357-361 (1977).
31. Tonkonogy, S. L. & Winn, H. J. *J. Immun.* **116**, 835-841 (1976).
32. Kimura, S., Tada, N., Nakayama, E., Liu, Y. & Hammerling, U. *Immunogenetics* **14**, 3-14 (1981).
33. Wall, K. A., Lorber, M. I., Loken, M. R., McClatchey, S. & Fitch, F. W. *J. Immun.* **131**, 1056-1064 (1983).
34. Macphail, S. & Stutman, O. *Eur. J. Immun.* **14**, 318-324 (1984).
35. Janeway, C. A. & Katz, M. E. *J. Immun.* **134**, 2057-2063 (1985).
36. Peck, A. B., Janeway, C. A. & Wigzell, H. *Nature* **266**, 840-842 (1977).
37. Molnar-Kimber, K. & Sprent, J. *J. exp. Med.* **151**, 407-417 (1980).
38. Lynch, D. H. *et al. J. Immun.* **134**, 2071-2078 (1985).
39. Katz, M. E. & Janeway, C. A. *J. Immun.* **134**, 2064-2070 (1985).
40. Kourilsky, P. & Claverie, J.-M. *Ann. Inst. Pasteur/Immun.* **137D**, 3-21 (1986).

41. Blackman, M. *et al. Cell* **47**, 349-357 (1986).
42. Macphail, S. *et al. J. Immun.* **135**, 2967-2974 (1985).
43. Festenstein, H. *Transplantation* **18**, 555-557 (1974).
44. Jones, B. & Janeway, C. A. *Immunogenetics* **16**, 243-255 (1982).
45. Cerottini, J.-C., Engers, H. D., MacDonald, H. R. & Brunner, K. T. *J. exp. Med.* **140**, 703-717 (1974).
46. MacDonald, H. R., Engers, H. D., Cerottini, J.-C. & Brunner, K. T. *J. exp. Med.* **140**, 718-730 (1974).
47. MacDonald, H. R., Blanc, C., Lees, R. K. & Sordat, B. *J. Immun.* **136**, 4337-4339 (1986).
48. Heitzmann, H. & Richards, R. M. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3537-3541 (1984).
49. Oi, V. T., Glazer, A. N. & Stryer, L. *J. Cell Biol.* **93**, 981-986 (1982).
50. MacDonald, H. R. & Zaeck, P. *Cytometry* **3**, 55-58 (1982).
51. Ceredig, R., Lowenthal, J. W., Nabholz, M. & MacDonald, H. R. *Nature* **314**, 98-100 (1985).
52. Nabholz, M. *et al. Immun. Rev.* **51**, 125-156 (1980).
53. Leo, O., Foo, M., Sachs, D. H., Samelson, L. E. & Bluestone, J. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1374-1378 (1987).
54. Landegren, U. *J. Immun. Meth.* **67**, 379-488 (1984).

LETTERS TO NATURE

Discovery of a binary millisecond pulsar in the globular cluster M4

A. G. Lyne*, J. D. Biggs*, A. Brinklow*,
M. Ashworth† & J. McKenna*

* University of Manchester, Nuffield Radio Astronomy Laboratories,
Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

† University of Manchester Regional Computer Centre, Oxford Road,
Manchester M13 9PL, UK

We report the discovery of an 11-ms pulsar, PSR1620-26, in the closest globular cluster, M4 (NGC6121). It is the fifth millisecond pulsar to be found, and the second in a globular cluster. Unlike the other cluster pulsar, PSR1821-24 in M28 (ref. 1), PSR1620-26 is in a low-mass binary system. This provides strong support for formation mechanisms in which an old neutron star is spun up to millisecond periods by accretion during an X-ray binary phase. The orbit's small eccentricity is likely to have significant implications for both the age of the millisecond pulsar and its surface magnetic field. We conclude that the pulsar's current companion is, in fact, the star which was responsible for the spin-up.

Millisecond pulsars are probably the direct descendants of low-mass X-ray binary systems (LMXBs)^{2-4,30} in which an old neutron star is spun up by accretion of material from a companion. Present statistics indicate that the probability of LMXB occurrence is a few orders of magnitude higher for stars in globular clusters than in the Galactic disk⁵. This is emphasized by the fact that globular clusters contain ~0.05% of the mass of the Galaxy, but almost 20% of the LMXBs⁶. Globular clusters make particularly suitable birthplaces for LMXBs because of their high stellar concentrations and attendant higher rate of binary formation⁷. On the basis of these arguments, it might be expected that millisecond pulsars will exhibit the same bias⁸. However, no large-scale searches for pulsed emission have, until now, exploited this possible association, mainly because of insufficient computing resources.

Although millisecond pulse searches were not possible at the time, Hamilton, Helfand and Becker⁹ used the Very Large Array (VLA) in a survey of 12 globular clusters, looking for unresolved sources which might be millisecond pulsars. Their only promising detection was a weak source deep within the core of M28. Subsequent work on this object revealed a steep spectrum and a high degree of linear polarization, both of which are characteristic of pulsar radiation¹⁰. Recently, with the aid of the 76-m Lovell telescope at Jodrell Bank and a Cray X-MP supercomputer, this object was found to be a pulsar with a period of 3 ms (ref. 1).

Encouraged by this result, we have searched for pulsed emission from a further 24 nearby, dense globular clusters using the Lovell telescope. Here we present preliminary details of the

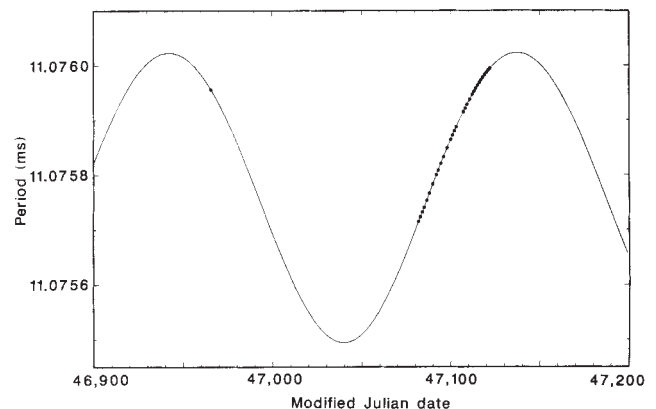


Fig. 1 Plot of pulse period as a function of epoch for PSR1620-26. The errors in the period measurements are typically 1% of the diameter of the symbols. Also shown is the period variation for the provisional binary orbit determined from the limited data span (see Table 1).

experiment which has led to the discovery of an 11-ms pulsar¹¹ in the globular cluster M4.

As part of this search programme, M4 was observed at 610 MHz on 19 June 1987. Total power outputs from 32 contiguous 125-kHz channels were recorded on magnetic tape every 300 μ s and, as with the observations of M28 (ref. 1), 1-bit digitization was used to reduce off-line storage requirements. The source was tracked for 78 min, during which about 14 million data samples were recorded from each of the 32 frequency channels.

The data were transferred to the Cyber 205 supercomputer at the University of Manchester Regional Computer Centre so that a search in periodicity and dispersion measure (DM) could be performed. The search algorithm first computed the two-dimensional fast Fourier transform of successive groups of 1

Table 1 Measured parameters of the PSR1620-26 system

Pulsar period	$11075.75 \pm 0.10 \mu\text{s}$
Dispersion measure	$63 \pm 2 \text{ cm}^{-3} \text{ pc}$
Right ascension (1950.0) (ref. 14)	$16 \text{ h } 20 \text{ min } 34.14 \text{ s} \pm 0.01$
Declination (1950.0) (ref. 14)	$-26^\circ 24' 58.0 \pm 0.2''$
Flux density (408 MHz)	$\sim 15 \text{ mJy}$
Projected semi-major axis, $a_p \sin i$	$64 \pm 7 \text{ light s}$
Orbital period, P_b	$195 \pm 20 \text{ days}$
Mass function, $f(m_p, m_c)$	$0.007 M_\odot$
Eccentricity, e	≤ 0.05
Assumed longitude of periastron, ω	0.0°
Time of periastron, T_0	$47,137 \pm 3 \text{ modified Julian date (MJD)}$
Projected orbital velocity	$\sim 7 \text{ km s}^{-1}$

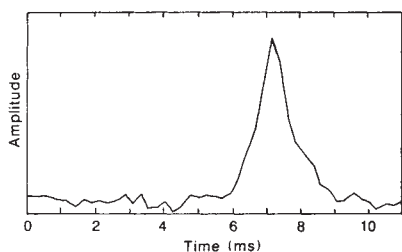


Fig. 2 The integrated pulse profile of PSR1620-26 at 408 MHz. The instrumental resolution is $\sim 300 \mu\text{s}$.

million samples from all 32 frequency channels and then output the pulse repetition frequency, DM and other relevant information of any prominent spectral features.

The output from the analysis of each sequence of 1 million data points was then inspected on the Starlink VAX 11/780 at Jodrell Bank, the 14 sets of output data being searched for coincidences in pulse frequency and DM. The best candidate had a peak signal-to-noise ratio of 9 at a frequency of 90.282 Hz, and appeared above the noise in 7 of the 14 analyses, with a DM in the range $40\text{--}68 \text{ cm}^{-3} \text{ pc}$. This frequency corresponds to a period at the Solar System barycentre of 11.076 ms.

On 12 October 1987, we observed M4 with the Lovell telescope at 408 MHz and confirmed the existence of the pulsar. We estimate that the distance to the pulsar is $2.2 \pm 0.8 \text{ kpc}$ using the observed value of DM, $63 \text{ cm}^{-3} \text{ pc}$, and the Lyne, Manchester and Taylor¹² electron distribution model. This estimate is compatible with the distance to the cluster, $\approx 2.0 \text{ kpc}$, determined from interstellar reddening studies of cluster main-sequence stars¹³. VLA observations at 1,450 MHz have subsequently shown there to be a highly polarized, unresolved 1.5 mJy source¹⁴ within the 1.2 arc min radius¹⁵ of the cluster core. The pulsar therefore almost certainly lies within the cluster core.

Our October observations also showed that the barycentric period was significantly shorter than it had been four months earlier, possibly owing to the Doppler effects of binary motion. Subsequent measurement of pulse arrival times have now confirmed that this pulsar is, indeed, in a binary system (see Fig. 1).

The characteristic properties of the new pulsar (designated PSR1620-26) are given in Table 1. The parameters of the fitted orbit must be regarded as provisional because of the limited time-span of the data. The fit also relies on an assumed value of zero for the period first derivative, $\dot{P}$. Other values of $\dot{P}$ do not degrade the quality of the fit, although the fitted orbital period and eccentricity do change slightly. Because of the long period of the binary system, about one year of timing data will be required before the period derivative can be determined with any precision. An integrated pulse profile obtained at 408 MHz is shown in Fig. 2. It consists of one moderately wide component with a duty cycle of $\sim 10\%$.

Some information about the masses of the two binary components is given by the pulsar's mass function, a property of only two observables, the projected semi-major axis, $a_p \sin i$, and orbital period, P_b :

$$f(m_p, m_c) = \frac{4\pi^2 (a_p \sin i)^3}{G P_b^2} = \frac{(m_c \sin i)^3}{(m_p + m_c)^2}$$

where i is the inclination of the orbit, and m_p and m_c are the masses of the pulsar and companion. Although these masses cannot be determined individually, m_c can nonetheless be estimated as a function of $\cos i$ for various neutron star masses, m_p . As shown in Fig. 3, for pulsar masses of $\sim 1 M_\odot$ and most orbital inclinations, the companion mass appears to lie in the range $0.3\text{--}0.4 M_\odot$.

The discovery of the millisecond pulsar PSR1821-24 in the globular cluster M28 (ref. 1), and now PSR1620-26 in M4 (ref. 11), provides strong support for the association of millisecond

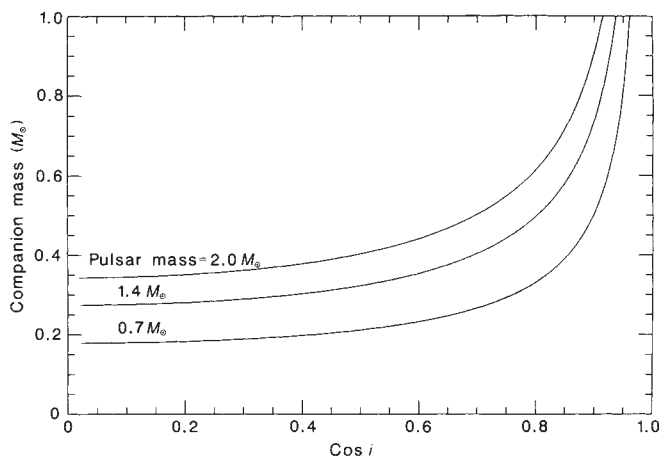


Fig. 3 The mass of the companion plotted as a function of $\cos i$ for various neutron star masses, using the value for the mass function given in Table 1.

pulsars and globular clusters. Moreover, the presence of PSR1620-26 in a binary system gives direct evidence that the formation of millisecond pulsars is related to the evolution of low-mass binary systems. Conventionally, millisecond pulsars are thought to arise within such binaries in the following manner: a massive ($\sim 1.3 M_\odot$) white dwarf or neutron star, left over from the early phase of stellar evolution in the cluster, tidally captures a main-sequence star during a close encounter⁷. The mass of the companion star at this time is likely to be $\leq 0.8 M_\odot$ because heavier stars would already have evolved away from the main sequence¹⁶. When this companion evolves into a giant, a period of sustained, conservative mass transfer will lead to copious X-ray production from the infalling material. The mass transfer will also spin up the compact object and make the orbit circular (ref. 17 and refs therein). If the primary is already a neutron star, then a millisecond pulsar will probably result after $10^7\text{--}10^8 \text{ yr}$. However, given the correct combination of accretion rate and initial mass, a white dwarf primary may accrete enough mass during the X-ray phase to undergo accretion-induced collapse to a neutron star without disrupting the system¹⁸. The millisecond pulsar may then be formed either as a result of further accretion and spin-up¹⁸ with accompanying recircularization of the orbit, or through conservation of angular momentum during the collapse of the white dwarf¹⁹. In the latter case, however, there may be a problem explaining the small observed eccentricity. The orbital period and the projected semi-major axis of PSR1620-26 are entirely within the ranges predicted by these formation scenarios.

Combining these ideas with studies of the late evolution of moderate mass giants, it has been suggested that there is a relationship between the final orbital period and the mass of the remnant dwarf star (ref. 31). The mass of the companion so derived lies within the $0.3\text{--}0.4 M_\odot$ range estimated from the pulsar mass function, adding some support to this proposal.

The 11.1-ms pulse period is rather longer than those of the four other millisecond pulsars known, which range from 1.6 to 6.1 ms. This has implications for the magnetic field and age of PSR1620-26, because both have a direct influence on the pulse period. The strength of the magnetic field determines both the minimum rotation period that can be achieved during spin-up and the subsequent rate at which the neutron star spins down. The neutron star will be spun-up by accretion until material at the inner edge of the accretion disk corotates with the magnetosphere. The minimum period that can be achieved is then given by the spin-up equation^{3,20}:

$$P_{\min} \approx 2.4 B_9^{6/7} (M/M_\odot)^{-5/7} (\dot{M}/\dot{M}_{\text{Edd}})^{-3/7} R_6^{18/7} \quad (1)$$

where B_9 is the magnetic field in units of 10^9 G , $\dot{M}_{\text{Edd}}$ is the

Eddington accretion limit, $\sim 10^{-8} M_{\odot} \text{ yr}^{-1}$, and R_6 is the neutron star radius in units of 10^6 cm . Neutron stars with weak magnetic fields should therefore be spun up to the shortest periods. Once accretion ceases, the pulsar will spin down by magnetic dipole radiation to a period given by the equation²¹

$$P^2(t_9) = P_{\min}^2 + 61.5 B_9^2 t_9 \quad (2)$$

where both periods are in ms, B_9 is assumed constant, t_9 is the spin-down age (the time elapsed since the end of accretion) in 10^9 yr , and standard values are used for the neutron star's radius and moment of inertia. The maximum value for the magnetic field from equation (1) is $\sim 8 \times 10^9 \text{ G}$ if the present observed period is close to the equilibrium spin-up value (that is, the millisecond pulsar is young). On the other hand, if the magnetic field is smaller, accretion rates up to the Eddington limit can spin up the neutron star to the minimum possible rotation period, $\sim 1 \text{ ms}$ (refs 22, 23). For the period then to decay to 11 ms within the lifetime of the globular cluster ($\sim 10^{10} \text{ yr}$), equation (2) implies $B \geq 4 \times 10^8 \text{ G}$. Magnetic fields within these two limits will result in intermediate values of spin-up period and spin-down age.

Of course, if the accretion is interrupted before the minimum spin-up period can be achieved, perhaps as the result of disturbances to the binary system by a stellar encounter or the accretion-induced collapse of the white dwarf primary, then the spin-down age will be less than that expected from equations (1) and (2).

In the future, a determination of the pulsar period derivative from continued timing observations will yield the magnetic field directly and hence allow estimates to be made of both the equilibrium spin-up period and the age of the millisecond pulsar. This will also provide useful information on the asymptotic behaviour of magnetic fields in old pulsars^{21,24}.

We may be able to deduce more about the history and age of PSR1620-26 because it is still in a binary system. Furthermore, the orbit is very nearly circular indicating that, despite the high density of stars within the core of the cluster ($\rho \approx 10^4 M_{\odot} \text{ pc}^{-3}$, ref. 15), the binary system has not suffered any major encounters with other stars that would either have made the orbit eccentric or disrupted the binary^{25,26}. The PSR1620-26 system may be prone to these effects because it is a rather wide binary (mean separation $\sim 1 \text{ AU}$) and the projected orbital velocity (7 km s^{-1}) is close to the cluster dispersion velocity of 5 km s^{-1} (ref. 15). An estimate of the age of a binary system in a globular cluster from its observed eccentricity (induced by glancing encounters) is best done by numerical simulation experiments. Those undertaken by Hut and Paczynski²⁷ for somewhat tighter systems suggest that in the core of M4 it would take $\leq 3 \times 10^9 \text{ yr}$ to produce the observed eccentricity. However, a more reliable estimate of the age of the PSR1620-26 binary awaits simulations based on its particular parameters.

The companion star's mass also has implications for the history of the binary system. At $\sim 0.3 M_{\odot}$, it is close to the cutoff mass⁵ below which all solitary stars evaporate from the cluster. Thus if a close encounter were to occur, resulting in the formation of an unstable triple system, the current companion is likely to be ejected because it would necessarily be the least massive star of the three^{28,29}. It would then, incidentally, retain most of its orbital velocity (30 km s^{-1}) and probably escape from the cluster^{25,26} ($v_{\text{esc}} \approx 20 \text{ km s}^{-1}$, ref. 15). We conclude, therefore, that the present companion is the star from which mass was transferred, creating the millisecond pulsar we now observe.

So far, two millisecond pulsars have been found in surveys of about 36 globular clusters undertaken in our present work and by Hamilton *et al.*⁹ If the luminosity function of millisecond pulsars is similar to that of the more numerous, slowly rotating, radio pulsars, a planned improvement in our system's sensitivity of about three should result in the discovery of a few more millisecond pulsars in the sample of 36 globular clusters. Of course, other millisecond pulsars may await discovery in the globular clusters not surveyed to date. We think it likely that

several new globular cluster millisecond pulsars will be found in the near future.

Received 30 November 1987; accepted 20 January 1988.

1. Lyne, A. G. *et al.* *Nature* **328**, 399-401 (1987).
2. Smarr, L. L. & Blandford, R. D. *Astrophys. J.* **207**, 574-588 (1976).
3. Radhakrishnan, V. & Srinivasan, G. in *Proc. 2nd Asian-Pacific Regional Meeting of the IAU* (eds Hidayat, B. & Feast, M. W.) 423-432 (Tira Pustaka, Jakarta, 1984).
4. Alpar, M. A., Cheng, A. F., Ruderman, M. A. & Shaham, J. *Nature* **300**, 728-730 (1982).
5. Verbunt, F. & Hut, P. in *IAU Symp. 125: The Origin and Evolution of Neutron Stars* (eds Helfand, D. J. & Huang, J. H.) 187-197 (Reidel, Dordrecht, 1987).
6. Verbunt, F. *Astrophys. J.* **312**, L23-L25 (1987).
7. Fabian, A. C., Pringle, J. E. & Rees, M. J. *Mon. Not. R. astr. Soc.* **172**, 15P-18P (1975).
8. Fabian, A. C., Pringle, J. E., Verbunt, F. & Wade, R. A. *Nature* **301**, 222-223 (1983).
9. Hamilton, T. T., Helfand, D. J. & Becker, R. H. *Astr. J.* **90**, 606-608 (1985).
10. Erickson, W. C., Mahoney, M. J., Becker, R. H. & Helfand, D. J. *Astrophys. J.* **314**, L45-L47 (1987).
11. Brinklow, A., Lyne, A. G., Biggs, J. D. & Ashworth, M. *IAU Circ. No.* 4470 (1987).
12. Lyne, A. G., Manchester, R. N. & Taylor, J. H. *Mon. Not. R. astr. Soc.* **213**, 613-639 (1985).
13. Alcaïno, G. & Liller, W. *Astrophys. J. Suppl. Ser.* **56**, 19-65 (1984).
14. Goss, W. M., Kulkarni, S. R. & Lyne, A. G. *Nature* **332**, 47-49 (1988).
15. Webbink, R. F. in *IAU Symp. 113: Dynamics of Star Clusters* (eds Goodman, J. & Hut, P.) 541-577 (Reidel, Dordrecht, 1985).
16. Grindlay, J. E. in *IAU Symp. 125: The Origin and Evolution of Neutron Stars* (eds Helfand, D. J. & Huang, J. H.) 173-185 (Reidel, Dordrecht, 1987).
17. van den Heuvel, E. P. J. in *Birth and Evolution of Neutron Stars: Issues Raised by Millisecond Pulsars* (eds Reynolds, S. P. & Stinebring, D. R.) 86-106 (National Radio Astronomy Observatory, Charlottesville, 1984).
18. Helfand, D. J., Ruderman, M. A. & Shaham, J. *Nature* **304**, 423-425 (1983).
19. Channugam, G. & Brecher, K. *Nature* **329**, 696-698 (1987).
20. van den Heuvel, E. P. J. *Ann. N.Y. Acad. Sci.* **302**, 14-35 (1977).
21. van den Heuvel, E. P. J., van Paradijs, J. A. & Taam, R. E. *Nature* **322**, 153-155 (1986).
22. Papaloizou, J. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **184**, 501-508 (1978).
23. Henrichs, H. F. & van den Heuvel, E. P. J. *Nature* **303**, 213-216 (1983).
24. Kulkarni, S. R. *Astrophys. J.* **306**, L85-L89 (1986).
25. Hills, J. G. *Astr. J.* **80**, 809-825 (1975).
26. Heggie, D. C. *Mon. Not. R. astr. Soc.* **173**, 729-787 (1975).
27. Hut, P. A. & Paczynski, B. *Astrophys. J.* **284**, 675-684 (1984).
28. Hills, J. G. *Astr. J.* **82**, 626-630 (1977).
29. Spitzer, L. & Mathieu, R. D. *Astrophys. J.* **241**, 618-636 (1980).
30. Webbink, R. F., Rappaport, S. & Savonije, G. J. *Astrophys. J.* **270**, 678-693 (1983).
31. Joss, P. C., Rappaport, S. & Lewis, W. *Astrophys. J.* **319**, 180-187 (1987).

Radio synthesis observations of the globular cluster M4

W. M. Goss*, Shrinivas R. Kulkarni† & A. G. Lyne‡

* National Radio Astronomy Observatory, Socorro, New Mexico 87801, USA

† Department of Astronomy, California Institute of Technology, Pasadena, California 91125, USA

‡ University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

The presence of an 11-ms pulsar in the direction of the globular cluster M4 and at about the same distance has recently been reported^{1,2}. Because of the large primary beam of the antenna used in the discovery, the relation of the pulsar to the cluster remained uncertain. We have made sensitive radio aperture synthesis observations and show that the pulsar indeed lies within the core of the globular cluster as projected onto the sky. We also report the absence of any other point sources within the core of the cluster having flux densities greater than one-seventh of that of the pulsar.

The globular cluster M4 was observed at 20 cm with the Very Large Array (VLA) of the National Radio Astronomy Observatory in three periods on 18 October, 19 October and 23 October 1987. The configuration of the VLA was the A/B array (the northern arm in the A array and the southwestern and southeastern arms in the B array) which produced a beam of 4.2 by 2.3 arc s, with the major axis at a position angle of 51°. The total observing time was 6.25 h and the observing frequencies were 1,465 and 1,515 MHz, each with a 50-MHz bandwidth. About 30 min of the observations of 23 October 1987 were contaminated by interference; these data were excised.

Following standard procedures we made a map of data from all the observations and bands. The final map had an r.m.s. of

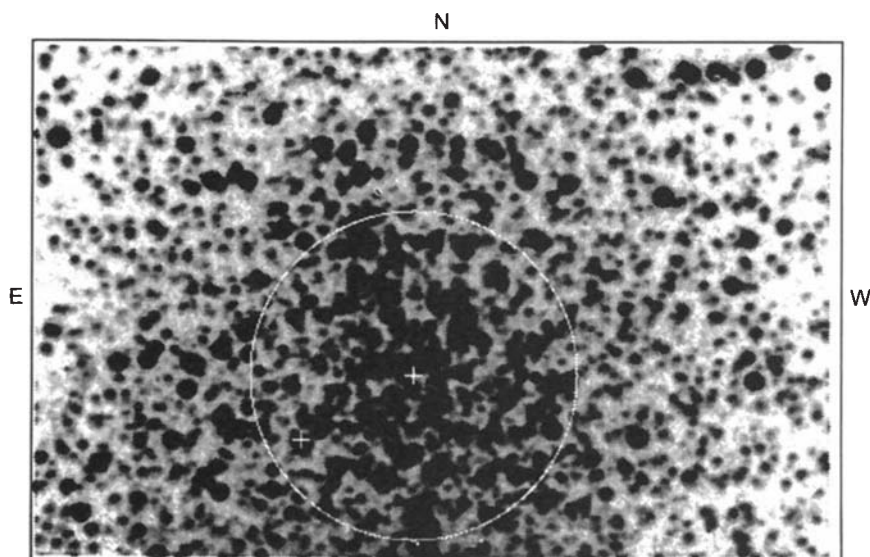


Fig. 1 A CCD image of the central region of the globular cluster M4. North is up and east to the left. A circle with a radius of 55 arc s marks the boundary of the cluster core. The centroid of the core is marked by a cross, as is the 11-ms pulsar, PSR1620-26, which is located 44 arc s to the south-east. The bright star close to the cross is an artefact. The overall radio-optical positional uncertainty is ~ 1.5 arc s.

0.060 mJy in total intensity. The delay beam of the VLA in this configuration limited the field of view to ~ 6 arc min (full width to half power). Two sources were detected: a 36-mJy source located 4 arc min to the north-west of the cluster core at a position of $\alpha(1950.0) = 16$ h 20 min 21.61 s, $\delta(1950.0) = -26^{\circ}21'19.7''$ and a $1.47 \pm 0.05(1\sigma)$ mJy source located only 44 arc s from the optical centroid of the globular cluster³ at $\alpha(1950.0) = 16$ h 20 min 34.14 s ± 0.01 s (1σ), $\delta(1950.0) = -26^{\circ}24'58.00'' \pm 0.15''(1\sigma)$. Polarization measurements obtained from the 23 October 1987 observations show that the former source has $<2\%$ linear polarization whereas the latter source showed a linear polarization of $31\% \pm 4\%$ (1σ).

The linear polarization of the second source at the frequency of the VLA observations is unprecedented in extragalactic sources, but is characteristic of the majority of pulsars. Furthermore, the flux density of the source is roughly that expected from the pulsar observations. Assuming a typical pulsar spectral index of -2 and the observed mean flux density of 15 mJy at 408 MHz (ref. 2), we estimate a mean flux density of 1.1 mJy at 1,490 MHz which is compatible with the observed flux density of 1.5 mJy. Taken together, these two observations clearly identify the second source as the 11-ms pulsar itself; hereafter we refer to this source as PSR1620-26. The source to the north-west of the core is not particularly distinctive and is most likely an extragalactic source.

In Fig. 1, we show an image of the central region of M4, obtained with an RCA CCD camera at the CTIO (Cerro Tololo Inter-American Observatory) 60-inch optical telescope by King and Djorgovski in the Johnson B-band. The intensity centroid was determined by the application of the mirror autocorrelation technique⁴. The size of the region used in the centroid analysis and the extent of smoothing of the CCD image lead to an overall uncertainty of 1.5 arc s (each axis) in the centroid. This is the principal source of error in our combined optical and radio astrometry. For the equatorial position of the optical centroid, we used the position quoted by Shawl and White³. From the known pixel scale of the CCD image (0.542 arc s per pixel) and the VLA position we were able to locate the position of the pulsar in the CCD image (shown by a cross). The boundary of the core as defined by the core radius is shown as a white circle. We used 55 arc s for the core radius (r_c) of the cluster; this value is a more recent determination by Djorgovski and King (personal communication) and is significantly lower than the standard value⁵ of 74 arc s.

The location of millisecond pulsars within the cores of the globular cluster M28 (ref. 6) and now M4 is in full accordance with theoretical expectation because neutron stars are sig-

nificantly more massive than the typical field star ($M \leq 0.5 M_{\odot}$). Indeed, the distribution of r_p/r_c can be used to determine statistically the mass of pulsars in units of the mass of the average cluster star; here r_p is the angular separation of the pulsar from the cluster core. Previous studies⁷ used low mass X-ray binaries (LMXBs) with relatively poor X-ray positions ($\sigma \approx 1.5$ arc s). In due course, pulsars such as PSR1620-26 with much better positions ($\sigma \approx 0.15$ arc s) can be used with great advantage to determine the mass of neutron stars with higher precision.

Examination of our radio image reveals the absence of any other source with flux density greater than 200 μ Jy. This can be compared with the current sensitivity of 1.5 mJy at 1,400 MHz of the Jodrell Bank globular cluster survey^{2,6}. Assuming a spectral index of -2 , this null detection implies that there are no other observable millisecond pulsars in the core of M4 with flux density greater than 2.3 mJy at 400 MHz. This comparison illustrates the high sensitivity of aperture synthesis observations. In addition, aperture synthesis observations are equally sensitive to fast pulsars in slow and rapid binary orbits whereas the pulsar search observations can be completely insensitive to fast pulsars in binary systems with small orbital periods. Of course, the principal disadvantage of aperture synthesis observations is that the mere detection of a point source does not mean that it is a pulsar. However, in the context of this discussion, aperture synthesis observations establish an upper limit to the number of pulsars above a certain flux density.

If millisecond pulsars are assumed to obey the same luminosity function as the standard (long period) pulsars, $n(L) \propto L^{-2}$ (ref. 8); the detection of PSR1620-26 and PSR1821-24 implies the presence of ≥ 12 other pulsars with flux densities greater than 200 μ Jy (at 1,400 MHz) within the core regions of the three dozen globular clusters searched in the course of the discovery of these two pulsars. Clearly, additional aperture synthesis observations of a large number of nearby dense clusters will be able to constrain the luminosity function of millisecond pulsars.

Following the discovery of a millisecond pulsar in M28 there was hope that observations would eventually constrain the many mechanisms that were proposed to explain the formation of millisecond pulsars in globular clusters^{9,10}. Unfortunately, the outcome of the most prolific mechanism, the collision of a neutron star with a main sequence star, is uncertain; possible final products include black holes, a slow period neutron star and a millisecond pulsar. Nevertheless, it may be that future observations, aperture synthesis or single dish pulse searches, will be able to improve the statistics sufficiently to distinguish between the different mechanisms.

S.R.K. would like to thank S. Djorgovski for immense help rendered in the analysis of the optical data, I. King for the CCD image of M4 and R. Romani for several suggestions that improved this paper. The National Radio Astronomy Observatory is operated by Associated Universities Inc., under contract with the National Science Foundation.

Received 30 November 1987; accepted 20 January 1988.

1. Brinklow, A., Lyne, A. G., Biggs, J. D. & Ashworth, M. *IAU Circ. No. 4470* (1987).
2. Lyne, A. G., Biggs, J. D., Brinklow, A., Ashworth, M. & McKenna, J. *Nature* **332**, 45-47 (1988).
3. Shawl, S. J. & White, R. E. *Astr. J.* **91**, 312-316 (1986).
4. Djorgovski, S. & King, I. R. *Astrophys. J.* **277**, L49-L52 (1984).
5. Webbink, R. F. in *IAU Symp. on Dynamics of Star Clusters* (eds Goodman, J. & Hut, P.), 541-577 (Reidel, Dordrecht, 1985).
6. Lyne, A. G. *et al. Nature* **328**, 399-401 (1987).
7. Grindlay, J. E., Hertz, P., Steiner, J. E., Murray, S. S. & Lightman, A. P. *Astrophys. J.* **282**, L13-L16 (1984).
8. Manchester, R. N. & Taylor, J. H. *Pulsars* (Freeman, San Francisco, 1977).
9. Romani, R. W., Kulkarni, S. R. & Blandford, R. D. *Nature* **329**, 309-310 (1987).
10. Verbunt, F., van den Heuvel, E. P. J., van Paradijs, J. & Rappaport, S. A. *Nature* **329**, 312-314 (1987).

Efficient detection of water masers

B. M. Lewis* & D. Engels†

* Arecibo Observatory, PO Box 995, Arecibo, Puerto Rico 00612, USA

† Hamburger Sternwarte, Gojenbergsweg 112, D-2050 Hamburg 80, FRG

During the past 15 years almost 190 natural sources of 22 GHz water maser emission have been found in the dusty circumstellar shells of late-type stars undergoing mass-loss. Numbered among these sources are ~75% of the Mira variables within 400 pc of the Sun¹, although only 34% of Miras in general are associated with OH masers². Stars enshrouded by circumstellar matter are easily recognized today by the characteristic colour-range of their far infrared flux³. The accuracy of this discrimination has been well-tested during searches for OH masers⁴⁻⁶, which unlike water masers⁷ depend on the infrared for their pumping. Colour-selecting IRAS (Infrared Astronomy Satellite) sources provides a convenient means of pinpointing likely new water masers, an approach first tested by Engels *et al.*⁴. We report here on the initial results of a more extensive survey, during which we found 60 masers, 49 of them detections. Most are associated with 1,612-MHz masers, but 7 are sources without other known masers. Our results rule out the binary hypothesis^{2,8} as a plausible general explanation for the existence of the many sources having far infrared colours typical of OH/infrared stars, which show no sign of any 1,612-MHz emission. Our data support the idea that sources evolve in both their colours and their rate of mass-loss.

Supergiant stars and stars on the asymptotic giant branch lose mass at rates ranging from 10^{-7} to more than $10^{-4} M_{\odot} \text{ yr}^{-1}$. This material cools as it moves away from the star, allowing both dust grains and molecules to form. As the dust-shell thickens, more of the stellar luminosity is absorbed and re-emitted as blackbody radiation in the far infrared at the lower temperature of the dust. Circumstellar material about evolved stars has an almost unique colour signature, which can rarely be confused with that from other objects. This led Habing⁹ to call such objects DGEs, or dusty gas enveloped stars. The DGEs include all Miras with circumstellar matter, and almost all water masers occur in DGEs. Searches⁴⁻⁶ for 1,612 MHz masers have already exploited this ability to select DGEs from the *IRAS Point Source Catalogue* (PSC)¹⁰, and refined the uncorrected colour range within which DGEs are known to be associated with OH masers. But ~50% of DGEs in general⁸, and 66% of Mira variables in particular², with much the same colours and galactic distribution, are not associated with 1,612-MHz masers. Existing compi-

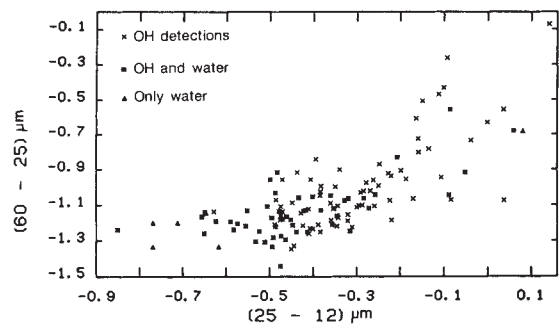


Fig. 1 Far infrared colour-colour diagram for the IRAS sources investigated in this survey for water maser emission. The colours are calculated as νS_{ν} from the fluxes listed in the *IRAS Point Source Catalogue*, assuming the adjustments appropriate to a 300 K blackbody.

lations¹¹ show that water masers occur in association with strong 1,612-MHz masers (type II), with 1,665/7-MHz masers (type I) and in sources without any OH emission. This is consistent with water masers arising in a shell closer to the exciting star than the OH in accord with their higher excitation energy ($T \approx 650$ K), with results^{12,13} from very long baseline interferometry (VLBI), and with OH being generated by the photo-dissociation of water by interstellar ultraviolet radiation. Water masers develop when the requisite shell number density is sufficient to excite them collisionally⁷: an $\dot{M} \approx 10^{-7} M_{\odot} \text{ yr}^{-1}$ suffices for Miras¹. Thicker shells are needed to support mainline OH masers, and even thicker ones for 1,612-MHz masers. Bowers and Hagen¹ show that this progression is accompanied by an increase in mass-loss. The occurrence of water masers in DGEs selected by infrared colour criteria should follow a similar pattern.

Our observations of the $6_{16}-5_{23}$ transition of water at 22 GHz, made with the 100-m radio-telescope at Effelsberg in June 1987, use a K-band maser as a front-end receiver. In clear weather this provides a 60-K system temperature at the zenith. Generally sources were observed for 5 min on source followed by 5 min at a comparison position. The signals were processed by two autocorrelation spectrometers to give a resolution of 0.2 km s^{-1} and a velocity coverage of $\pm 80 \text{ km s}^{-1}$. A lower resolution was used in search mode. Complete details of the observational technique and reduction methods are given by Engels *et al.*¹⁴.

Our source list is identical with the selection of DGEs examined in the initial stage of the Arecibo survey for 1,612-MHz masers⁶, for which the masing properties at 1,665/7 MHz are known (B. M. L., J. Eder and Y. Terzian, in preparation), as well as the emission velocity for objects with OH masers. We chose to observe the sources with the strongest 1,612-MHz fluxes, and the sources with $(60-25) \mu\text{m} > -0.9$. The overall distribution of objects on the sky is such that we could also observe all sources in the sample with right ascensions smaller than 19 h and $b^{\text{II}} > 10^{\circ}$. From a total of 177 DGEs examined, 61 have water masers. Of the 12 already known sources, all but 19493 + 2905 were redetected. Our detection rate for new masers in Table 1 is 49/166 or 29.5%. This is much better than recent searches without infrared data, where rates of 8.8% (ref. 15) and 3.3% (ref. 16) occur. Our success rate of 45/115 or 39% for DGEs with known 1,612-MHz masers rises to 17/25 or 68% for those with $|b^{\text{II}}| > 10^{\circ}$, where there are no sensitivity limitations on the water detection rate. Only 17% of similar DGEs without 1,612-MHz masers have water masers however. Although working with DGEs enhances the detection rate, it is clear that whatever conditions favour the occurrence of a 1,612-MHz maser, also favour the presence of a water maser.

Figure 1 shows that most of our water masers are associated with 1,612-MHz type II masers. Some of these exhibit weak mainline activity. Few infrared-selected DGEs are type I

Table 1 IRAS DGEs with water masers

Source	Position		Colour indices		(V) km s	T (K)	Nom comp	Source	Position		Colour indices		(V) km s	T (K)	Nom comp		
LII	BII	(25-12)	(60-25)	LII				BII	(25-12)	(60-25)							
14219+2555*	34†	69	-0.71	-1.27	3	0.2	1	19061+1041	44	1	-0.36	-1.05	29	1.2	2		
15060+0947	11	53	-0.52	-1.25	67/5	-8	1.1	3	19067+0811	42†	0	+0.06	-0.68	77	0.2	1	
15262+0400	8	46	-0.58	-1.24	67/5	45	0.6	1	19069+0916	43	0	-0.09	-0.56	27	3.5	5	
15298+0348*	9†	45	-0.77	-1.33		21	4.7	1	19071+0946	44†	1	-0.05	-0.92	67/5	11	0.6	1
16211+3057*	51	44	-0.65	-1.15		39	0.3	1	19087+0323	38	-3	-0.28	-1.06	67	10	1.2	2
16235+1900	35†	40	-0.85	-1.24	67/5	-17	5.2	2	19088+1129	45	1	-0.48	-1.44	67/5	31	0.5	2
16260+3454	56	44	-0.51	-1.11	67	59	1.0	3	19116+1511	49	2	-0.48	-0.92	67	70	0.8	2
16560+2252	43	34	-0.65	-1.14	67	-5	2.2	1	19134+2131*	55	5	+0.08	-0.68	-66	0.27	5	
17050+1714	38	31	-0.62	-1.19		-77	13	1	19161+2343	57	5	-0.32	-1.07	67	15	0.9	1
17123+1107*	32†	26	-0.77	-1.20		9	25	1	19186+0315	39	-5	-0.32	-1.25	-34	8.5	2	
17230+0113	24	20	-0.47	-1.16	67/5	-25	17	7	19288+2923	63	5	-0.26	-1.04	67	-23	1.7	1
17256+0504	28†	21	-0.49	-1.03	67	26	4.5	3	19303+1553	52	-2	-0.44	-1.25	-6	1.2	1	
17528+1144	37	18	-0.59	-1.19	67/5	89	0.06	1	19344+0016	39	-10	-0.35	-1.16	132	2.7	2	
17579+2335	49	21	-0.65	-1.26	67/5	3	1.5	1	19386+1513	52†	-4	-0.46	-1.29	17	0.5	1	
17594+0826*	35	15	-0.62	-1.33		16	0.3	1	19456+1927	57	-3	-0.33	-1.07	-78	0.8	1	
18069+0911	36†	14	-0.45	-1.18	67	-26	1.1	1	19462+2232	59†	-2	-0.47	-1.18	67	4	10	1
18076+3445	61†	23	-0.48	-1.14	67/5	14	5.5	2	19495+0835	48	-9	-0.49	-1.17	67/5	43	6.2	2
18231+0855	38	10	-0.55	-1.13	67/5	-10	1.3	2	19499+2141	59	-3	-0.42	-1.13	0	2.0	1	
18270+0326	34	6	-0.57	-1.20	67/5	-36	1.5	5	19508+2659	64	0	-0.38	-1.13	67	4	5.5	3
18375+0510	36	5	-0.51	-1.31	67	68	0.9	2	19522+1935	58	-4	-0.46	-1.17	67	55	1.2	1
18441+2216	52	11	-0.50	-0.96	67	-11	2.1	1	20097+1107	53	-12	-0.40	-1.05	-27	1.7	2	
18520+1014	42	4	-0.41	-1.25		32	0.5	3	20137+2838	68	-3	-0.09	-1.05	67	-67	0.8	2
18535+0726	40	2	-0.36	-1.21	67	46	0.1	2	20171+2732	67	-5	-0.43	-1.06	52	0.6	1	
18540+0302	36	0	-0.21	-0.83		85	0.1	1	20246+2813	69	-6	-0.47	-1.27	14	1.0	4	
18549+0905	42	3	-0.50	-1.17	67/5	16	8	3	20363+3401	75	-4	-0.49	-1.28	67/5	14	5.6	4
18556+0811	41	2	-0.49	-1.33	67/5	38	0.9	4	20482+3325	76	-7	-0.35	-1.12	43	0.6	4	
18567+0003	34	-2	-0.45	-1.21		47	0.2	3	21120+0736	59	-27	-0.53	-1.30	67	25	0.5	1
19017+0412	38	-1	-0.27	-1.12	67	98	3	1	21174+1747	68	-22	-0.49	-1.22	67/5	-79	2.3	2
19037+0824*	42	1	+0.24	-0.27		7	0.04	4	22402+1045	79	-41	-0.56	-1.22	67	-105	1.6	1
19043+1009	44	1	-0.42	-1.13	67	35	0.25	1	22516+0838	81	-44	-0.66	-1.16	67/5	3	1.7	2

(V) is the mean local standard of rest velocity. T(K) is the temperature of the strongest component, and 'Nom comp.' represents the number of discrete spectral components in the source.

* No 1,612-MHz maser.

† Previously known source of water.

sources with the strongest OH masers in one of the 1,665/7 MHz lines (B.M.L., J. Eder and Y. Terzian, in preparation). None of our water detections is associated with a type I source. This feature is probably linked with the larger expansion velocities typifying DGEs, which enables them to develop a thicker shell more quickly than the Miras. But seven objects exhibit water masers without associated OH emission: these generally have $|b^{II}| > 10^\circ$ and properties similar to Miras.

Herman and Habing² considered that those Miras lacking OH masers might well be members of binary systems. Their rationale is based on the fact that 1,612-MHz masers are both radiatively pumped and have long paths in the expanding shell. An orbiting companion could perhaps impose a large enough velocity gradient on the masing path to suppress the maser. When generalized to all DGEs, this hypothesis has the attractive feature of suggesting a Galaxy-wide means of estimating the frequency of single stars⁸. It is directly tested by our dataset, because Cooke and Elizur's⁷ models show that water masers are collisionally pumped, and so are unaffected by the multiplicity of the system. One then expects the presence of a water maser in a DGE to be independent of the presence of a 1,612-MHz maser, but about twice as frequent as there is a ~50% frequency of widely spaced binaries¹⁷ among Wolf-Rayet stars, O-type stars and bright main-sequence stars. On the other hand, our data show the opposite result: DGEs with 1,612 MHz masers are much more likely to exhibit water masers. This is most tellingly underlined by the subsamples of 25 objects with $|b^{II}| > 10^\circ$, for which the percentage of Type II masers with water to the percentage of other DGEs with water is 68:17. The explanation for DGEs without OH masers, or of those without water masers, must be sought elsewhere.

The key to understanding the varying occurrence of different masers in circumstellar envelopes appears to be in an evolution of their mass-loss rate. Bowers and Hagen¹ were able to relate the intensity of both the water and OH emission to the rate of

mass-loss. This trend was shown by Engels *et al.*¹⁴ to carry over to the OH/infrared stars, and to be directly related to their (25-12) μm colour index. Rowan-Robinson *et al.*¹⁸ find a similar correspondence with (25-12) μm colour while simulating the detailed shape of the far infrared spectrum obtained with various mass-loss rates, as the optical flux is reprocessed by dust. The additional feature suggested by our data is an absolute weakening of the water emission for the highest mass-loss rates (more positive values of the (25-12) μm index). The steep initial strengthening of water emission with $\dot{M}$ for $\dot{M} < 5 \times 10^{-6} M_\odot \text{yr}^{-1}$ is due to an increase in the radial zone capable of supporting masers. But collisional de-excitation at the inner margin of this zone noticeably weakens the intensity for $\dot{M} \geq 10^{-5} M_\odot \text{yr}^{-1}$, whereas our results indicate that the maser can be entirely quenched at yet larger $\dot{M}$.

OH/infrared stars are generally confined to the galactic plane³: our subgroup with $|b^{II}| > 10^\circ$ is therefore intermediate in properties between OH/infrared stars and Miras. Where Miras in the Solar Neighbourhood follow a trend from objects without water or OH ($\Omega = (25-12) \mu\text{m} = -0.84$), to objects with just water ($\Omega = -0.83$) to objects with both water and OH ($\Omega = -0.78$) which is closely tied to their mass-loss rate^{14,18}, optical selection of Miras does not generate a sample with OH but no water. This trend is paralleled among the DGEs with an offset: objects with no water or OH have $\Omega = -0.72$, objects with water and no OH have $\Omega = -0.70$, objects with water and OH have $\Omega = -0.56$, while objects with OH but no water have $\Omega = -0.35$. Evidently, increasing mass-loss eventually suffices to extinguish, or at least to greatly diminish, the intensity of water masers, a possibility already implicit in the decreasing ratio of water to OH emission with increasing mass-loss¹, and implied by the absence of water masers in the supergiant IRC10420. This trend is also supported by averages over the remaining DGEs with $|b^{II}| < 10^\circ$, where objects with OH and no water have $\Omega = -0.26$, whereas objects with OH and water have $\Omega =$

$\langle 25-12 \rangle \mu\text{m} = -0.37$, despite the sensitivity of water detections to variability and distance. We expect OH/infrared stars to lose their water masers, once their mass-loss rate is high enough to quench the masers in the exciting temperature range.

Seventy-five per cent of Miras have water masers, but only 41% of our group of comparably close DGEs do. Repeated water observations may slightly improve this percentage as water masers vary by large factors over short time intervals¹⁴, but re-observation is not likely to produce the water detection rate found among the Miras. In this, the subsample reverses the difference found between Miras (34%) and DGEs (46%) in their association with OH masers.

Cooke and Elitzur⁷, in modelling the pumping of water masers, find that the innermost masing radius is set by the number density at which quenching or collisional de-excitation dominates. The relatively large frequency of water masers among Miras would then be explained by the fact that Miras have lower mass-loss rates than OH/infrared stars, and so support water masing over a larger radial range. As the mass-loss processes increase with time, shells get thicker, Miras become obscured by dust and so may only be evident in the infrared, while the water masers get squeezed out to larger radii. Eventually the water masers may be quenched over the whole of their potential range. This appears to have happened with the supergiant star IRC10420, which has strong OH masers at 1,612, 1,665 and 1,667 MHz, but no hint of any water emission, and in 18095+2704, which has its strongest masers at 1,665/7 MHz. But, as

Table 1 shows, DGEs with both 1,612 MHz and mainline masers generally support water masers. These objects have a marginally smaller $\langle 25-12 \rangle \mu\text{m} \approx -0.39$ than the $\langle 25-12 \rangle \mu\text{m} = -0.33$ characterizing objects with water and OH, but no mainline masers. Mainline masers are usually restricted to the objects with the lower mass-loss rates.

This work is supported by the National Astronomy and Ionosphere Center, which is operated by Cornell University under a management agreement with the National Science Foundation.

Received 8 October; accepted 1 December 1987.

1. Bowers, P. F. & Hagen, W. *Astrophys. J.* **285**, 637-647 (1984).
2. Herman, J. & Habing, H. J. *Phys. Rep.* **124**, 255-314 (1985).
3. Lewis, B. M. in *Late Stages of Stellar Evolution* (eds Kwok, S. & Pottasch, S. R.) 75-78 (Reidel, Dordrecht, 1987).
4. Engels, D., Habing, H. J., Olmon, F. M., Schmid-Burgk, J. & Walmsley, C. M. *Astr. Astrophys.* **140**, L9-L12 (1984).
5. Lewis, B. M., Eder, J. & Terzian, Y. *Nature* **313**, 200-202 (1985).
6. Eder, J., Lewis, B. M. & Terzian, Y. *Astrophys. J. Suppl.* (in the press).
7. Cooke, B. & Elitzur, M. *Astrophys. J.* **295**, 175-182 (1985).
8. Lewis, B. M., Eder, J. & Terzian, Y. *Astr. J.* **94**, 1025-1034 (1987).
9. Habing, H. J. in *IAU Symp. 122: Circumstellar Matter* (eds Appenzeller, I. & Jordan, C.) 197-213 (Reidel, Dordrecht, 1987).
10. *IRAS Point Source Catalogue* (US Government Printing Office, Washington DC, 1985).
11. Engels, D. *Astr. Astrophys. Suppl.* **36**, 337-345 (1979).
12. Chapman, J. M. & Cohen, R. J. *Mon. Not. R. astr. Soc.* **220**, 513-528 (1986).
13. Diamond, P. J. et al. *Astr. Astrophys.* **174**, 95-102 (1987).
14. Engels, D., Schmid-Burgk, J. & Walmsley, C. M. *Astr. Astrophys.* **167**, 129-144 (1986).
15. Crocker, D. A., Hagen, W. *Astr. Astrophys. Suppl.* **54**, 405-416 (1983).
16. Dickinson, D. F. & Dinger, A. St. C. *Astrophys. J.* **254**, 136-142 (1982).
17. Abt, H. A. *Rev. Astr. Astrophys.* **21**, 343-372 (1983).
18. Rowan-Robinson, M., Lock, T. D., Walker, D. W. & Harris, S. *Mon. Not. R. astr. Soc.* **222**, 273-286 (1986).

Intensity profiles of dust near extended sources on comet Halley

N. Thomas*, D. C. Boice†, W. F. Huebner† & H. U. Keller*

* Max-Planck-Institut für Aeronomie, 3411 Katlenburg-Lindau, FRG
† Southwest Research Institute, San Antonio, Texas 78284, USA

Images obtained by the cameras on board the spacecraft encountering comet Halley in March 1986 showed that the dust activity was very non-uniform. Very fine structures (filaments) were observed as well as strong jet-like features originating from active areas¹. The dust particle density at large distances ($R > 100$ km) from the nucleus decreases proportional to $1/R^2$, that is, the intensity decreases proportional to $1/R$ as expected for free outflow. Flattening of the profile is evident within several tens of kilometres of the nucleus. As we show here, this can be explained by a model that considers the extended size and non-uniformity of the active region on the surface. A critical scale length, defined by the opening angle of a jet-like feature and the size of the source, can be introduced to describe the flattening. The model is consistent with the observations and provides the basis for studying other mechanisms such as particle fragmentation.

Halley Multicolour Camera (HMC) images of comet Halley show jet-like features and filaments which, in some cases, can be traced to the surface of the nucleus. The dust column density at large distances clearly exhibits $1/R$ dependence², as can also be demonstrated by plotting intensity I multiplied by the radial distance R (see Fig. 1). The quantity $P = IR$ is constant and a measure of the source strength of a point source, once the dust outflow is decoupled from the gas. The plot in Fig. 1 refers to an image taken 285,000 km from the comet in which the nucleus is essentially a point source less than three pixels in size and the origin for the R co-ordinate. However, at much smaller distances, the nucleus presents a finite size and the determination of the origin for R becomes less certain. A hypothetical origin for a point source can be determined from the best least squares fit to P using the angle of the inferred dust cone as a constraint

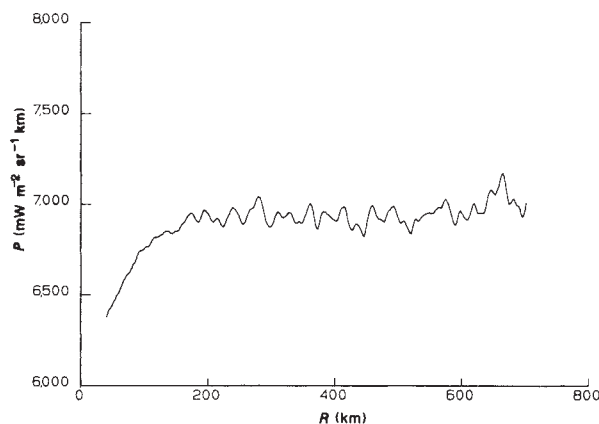


Fig. 1 A plot of $(1/2\pi) \int_0^{2\pi} IR d\theta$ (the azimuthal average) against R , the distance of the line of sight from the nucleus. The function is essentially constant at large distances, indicating free outflow of dust. Image 2468 was taken by HMC at a distance of 285,000 km.

on the uncertainty³. The intersection of the dust cone with the nucleus defines the size of the active area. This idealization forms the basis for a first approximation of the dust distribution that is valid at large distances and has been used for the initial interpretation of the Giotto and Vega images^{3,4}.

Difficulties with the point source model have been recognized in the analysis of images resolving the environment near the nucleus. If dust were emitted perpendicular to a smooth, spherical cap (which is equivalent to a point source) then the intensity profile would follow $1/R$ all the way to the surface. However, close to the nucleus flattening from the $1/R$ profile becomes apparent. This is clearly visible in Fig. 2. A satisfactory explanation of this phenomenon has not appeared^{3,4}. We present a qualitative argument supporting the effects of an extended source (see item (5) below) as the major influence on the intensity profile in the near-nucleus region.

Deviations of I from $1/R$ can be caused by several different

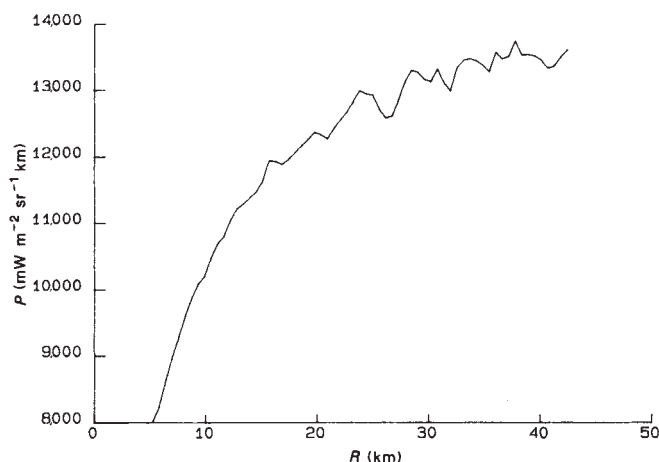


Fig. 2 A plot of $P = IR$ from image 3416, taken from 25,600 km, along the line of filament N3 (ref. 1). The curve indicates deviation from $1/R$ behaviour close to the nucleus.

mechanisms: (1) Radial acceleration of the dust by entrainment with the gas causes a steepening of the profile relative to $1/R$. This has not been observed because the small scale length for acceleration is less than or comparable to the overlapping of the intensity profiles of the dust and the active surface area. (2) Similarly, transverse flow of dust⁵ causes a reduction of P with increasing distance because of the decrease of dust particles remaining in the radial flow. (3) Since the optical thickness of the dust is small^{5,6}, the contribution to the flattening of the profile is insufficient. (4) Fragmentation of the dust grains can cause deviations of I from $1/R$ in either direction. Breaking off small particles that are invisible will decrease the cross section of the main component, thereby reducing P . On the other hand, fragmentation that leads to an increase in the number of visible particles increases P . (5) Effects caused by the extended size of the source area near the surface flatten the $1/R$ profile. Only the dust from a fraction of an active region contributes to the intensity at points in the line of sight just above the region of activity. The limited contribution of the source strongly reduces the P profile and predominates over all the other effects. This will be discussed in more detail below.

The first three of the above points are inconsequential. Item (4) requires some explanation. Combined with the assumption of the hypothetical point source, an intensity increase of about 25% (ref. 3) requires that almost every dust particle has to split within a few tens of kilometres above the nucleus. Only few fragmentations can occur thereafter. This scenario is not valid for icy grains or ice-mantle grains but is consistent for particles bonded by ice. However, the interpretation based on the hypothetical point source also requires that the surface be smooth and homogeneous to accommodate laminar and radial flow. HMC images⁷ show surface roughness at the limit of resolution (100 m) which is small compared to an observed active region (diameter 3 km). We therefore conclude that a model based on the hypothetical source point and fragmentation is physically inadequate.

The extended-size model resolves the above difficulties. It considers divergent flows from different parts of the surface under the assumption that these flows are tenuous enough not to influence each other. A schematic diagram of the geometry of the model is shown in Fig. 3 representing a jet-like feature with opening angle, α . The critical scale length is defined as that distance where the same opening angle α of an inward-

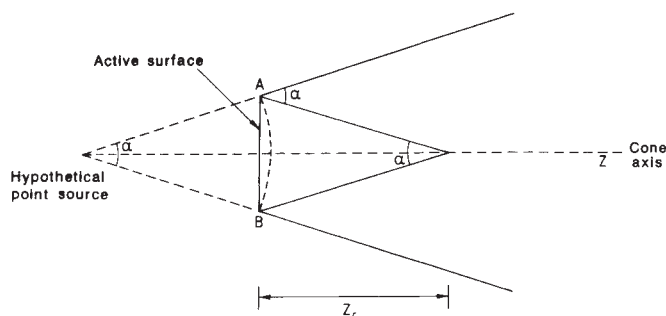


Fig. 3 At large distances ($\gg Z_c$) a dust jet has an opening angle, α , and constant source strength P ($R \approx Z$). Its origin appears to be at the hypothetical point source. At distances less than Z_c , emissions from points A and B no longer contribute to the dust density along the cone axis. Consequently a flattening of the intensity profile relative to $1/R$ is observed.

looking cone subtends the entire active area. (This is equal to the distance of the hypothetical point source from the active surface.) The point source approximation will be approached asymptotically at large distances. Within a few critical scale lengths, the deviations from the point source model are reduced to a few per cent. In the limit of $z \rightarrow 0$ the fraction of the active area subtended by the inward cone also approaches zero, and the column density viewed parallel to the surface leads to the limit $Iz \rightarrow 0$. This demonstrates that the flattening of the observed profile is an inherent feature of this model.

The physical description indicates that the active area emits dust into a cone with angular opening described by its associated jet-like feature. This divergent emission is related to the physical properties of the surface, while the hypothetical point source refers only to an idealized spherical surface. Therefore, we conclude from these observations that the surface is not smooth or homogeneous but must be characterized by varying curvature, roughness, and source strength to reflect the observed dust distribution. The extended source model forms the basis from which we can infer surface properties such as curvature and roughness by studying the details of the jet-like features. Properties related to the dynamics of the dust flow are refinements which can be incorporated into the model. Hydrodynamic interaction of neighbouring flows will occur if their densities are high and will result in less divergence. However, the fact that Iz can be very small near the surface indicates that taking into account the stereo geometry naturally explains the flattening of the observed profile of the dust distribution.

The extended source model encompasses the qualitative features of the observed dust emission far from and near to the nucleus and the properties of the nuclear surface, such as the surface topography and source strength and distribution. A mathematical model simulating the activity from an extended source by a distribution of point sources, each emitting dust into a narrow cone, needs to be developed.

This work was supported in part by funds of BMFT (FRG) and from the NASA Planetary Atmospheres Program.

Received 12 October; accepted 27 November 1987.

1. Thomas, N. & Keller, H. U. *Astr. Astrophys.* **187**, 843-846 (1987).
2. Keller, H. U. *et al. Nature* **321**, 320-326 (1986).
3. Thomas, N. & Keller, H. U. *Eur. Space Ag. spec. Publ.* **278**, 337-342 (1987).
4. Toth, I., Szegő, K. & Kondor, A. *Eur. Space Ag. spec. Publ.* **278**, 343-347 (1987).
5. Keller, H. U. *et al. Astr. Astrophys.* **187**, 807-823 (1987).
6. Sagdeev, R. Z. *et al. Eur. Space Ag. spec. Publ.* **250**, 11, 317-326 (1986).
7. Keller *et al. Nature* **331**, 227-231 (1988).

Extremely low N₂O concentrations in the springtime stratosphere at McMurdo Station, Antarctica

A. Parrish*, R. L. de Zafra, M. Jaramillo, B. Connor, P. M. Solomon & J. W. Barrett

State University of New York, Stony Brook, New York 11794, USA *and Millitech Corp., South Deerfield, Massachusetts 01373, USA

Nitrous oxide is an important tracer of transport, as well as the ultimate source gas for NO_x in the stratosphere. Its calculated chemical lifetime in the troposphere and lower stratosphere is longer than transport timescales worldwide, so that its mixing ratio is not expected to be strongly dependent on latitude or season. However, we have now made frequent measurements of stratospheric N₂O using the Stony Brook millimetre-wave remote sensing spectrometer at McMurdo Station, Antarctica, as part of the investigation of the 'ozone hole'¹ and find N₂O mixing ratios that are less than 1/5 at 20 km, and less than 1/10 at 25 km compared to values measured during the Antarctic summer. The observed mixing ratios are also much less than those predicted by global-scale models of stratospheric chemistry and dynamics. As the N₂O signal remained very weak when McMurdo was at the edges of the ozone hole and showed no signs of recovering during October, we conclude that the geographical and temporal extent of the region of low N₂O is comparable to, or greater than that of the ozone hole. Our results argue against theories that require springtime upwelling to explain the Antarctic ozone hole, and suggest that the air in the Antarctic lower stratosphere during late winter and early spring has been subjected to considerable downward transport.

The calculated chemical lifetime of N₂O in the troposphere is in the range of 100–200 years². Only in the stratosphere, above 15 km, does the lifetime become short enough to cause a significant latitudinal gradient. At 25 km, the measured mixing ratio in high-latitude summers can be one third to one half of that near the equator^{3–5}. Such latitudinal variations are also predicted by various theoretical models^{6–8}, which fall into two classes: those which predict a middle stratospheric gradient of about a factor of two^{5,6} between the equator and 80° S, and those which predict a gradient of an order of magnitude^{7,8}. Data available for comparison with these models include stratospheric and mesospheric sounder (SAMS) measurements⁵ for the latitude range 70° N to 50° S, a set of seven balloon-borne, grab-sampling measurements^{3,9} conducted in Antarctica during the austral summers of 1976 to 1979, and one set of infrared measurements¹⁰. These measurements are in better agreement with the models^{5,6} that predict the smaller latitudinal gradient. However, the results we report below show that there is, during the late austral winter and early spring, far less N₂O in the Antarctic middle stratosphere than present in the summer—and far less than observed at any other location. These new results are not predicted by any theoretical models.

We observe the $J = 11$ to 10 rotational transition of N₂O in emission at 276 GHz using an instrument consisting of a millimetre-wave heterodyne receiver and a multi-channel filterbank spectrometer having a spectral bandpass of 256 MHz and a resolution of 1 MHz. This instrument, which is described in detail elsewhere¹¹, has a calibration accuracy of ~12% and sufficient spectral coverage and resolution to obtain the altitude distribution of an observed species in the middle stratosphere by deconvolution of its pressure-broadened emission line-shape. The instrument does not, however, have sufficient spectral coverage to contain the full N₂O spectral line and hence provides only limited quantitative information at altitudes below 20 km.

The far wings of ozone lines bracketing our spectral window produce a weak concave background curvature. The magnitude

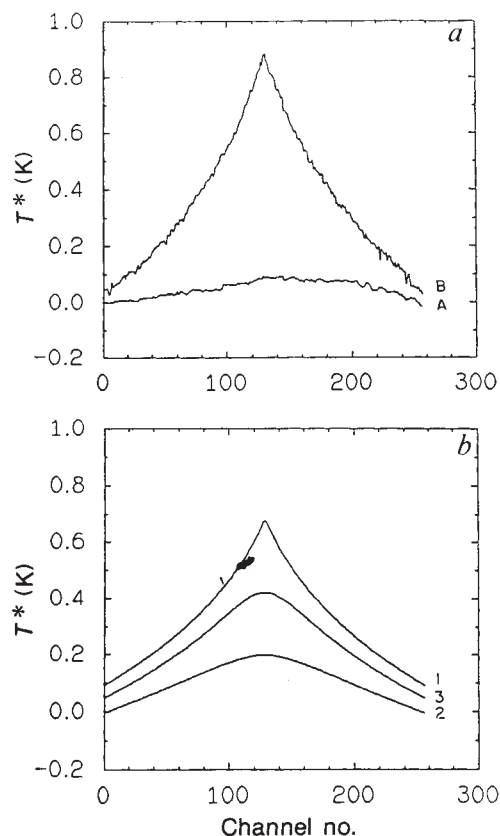


Fig. 1 Panel a: Spectrum A shows the average of data taken at McMurdo Station, Antarctica, from 12 September to 28 October 1986, excluding data from 23 and 24 September. Spectrum B shows data taken on 23 May 1986, at Mauna Kea, Hawaii¹², shown for comparison. Panel b: Curve 1 is a synthesized emission line spectrum corresponding to the altitude distribution predicted by the model of Harwood and Pyle, as reported in ref. 5 at 78° S; Curve 2 is a synthesized spectrum from the distribution of Solomon and Garcia (ref. 8 and private communication), also at 78° S. Curve 3 is a synthesized emission line corresponding to balloon-borne measurements made in the Antarctic summers of 1976 to 1979. This spectrum slightly underestimates the peak line intensity because data are only available to the balloon altitude limit of 30 km. The altitude distributions corresponding to spectra 1–3 are shown in Fig. 3. Abscissa is spectrometer channel number, at 1 MHz per channel. The N₂O line centre frequency is at channel number 128. Ordinate is intensity in temperature units referred to emission from the black body calibration standards.

of this curvature is small compared to the intensity of the N₂O signal expected from the Antarctic summertime N₂O profile; it affects our results significantly only because of the weakness of any observable signal from springtime N₂O. We have computed spectral backgrounds using specific balloon-sonde measurements of ozone, temperature and pressure made at McMurdo by Hofmann *et al.*¹² and find that the centres of these broad curves are from 0.065 to 0.11 K deep compared to the edges, with an average depth of 0.075 K (we express intensities in temperature units based on emission from black-body calibration standards). We have subtracted the computed ozone background appropriate for the daily conditions, as well as a small slope, from each spectrum shown in Figs 1a and 2. There are no other significant stratospheric background contributions in this spectral window.

The data shown in Figs 1 and 2 do contain some influence, probably instrumental, other than the N₂O and ozone background contributions; this is evident from the small (compared to expected Antarctic signal intensities) residual asymmetry seen in all scans in Fig. 2. Our conclusions regarding N₂O mixing ratios in the middle stratosphere are based on the fact that we

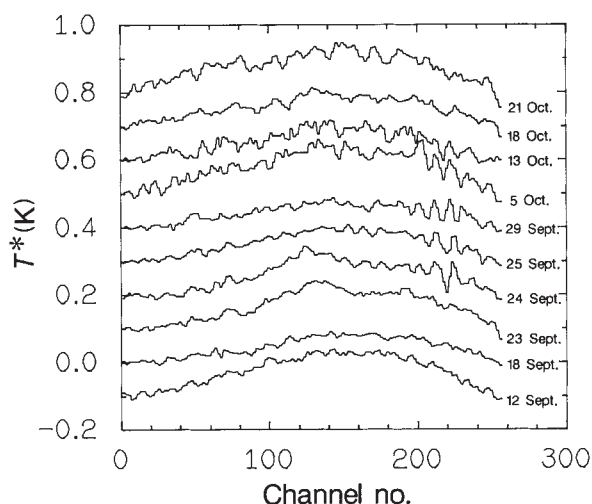


Fig. 2 Representative spectra taken at the N_2O emission line frequency. Each scan is offset from the preceding one by 0.1 K. Numbers to the right of the spectra give the day and month the data were taken. A calculation of the spectral background from nearby ozone lines has been subtracted from each spectrum, as has a linear baseline. The abscissa and ordinate are the same as in Fig. 1.

do not see well-defined central spectral features like those shown in Fig. 1; such features are easily distinguished from the broad curves which are characteristic of instrumental baselines. The N_2O spectral contribution from the stratosphere below about 20 km does resemble baseline curvature, and no certain conclusions can be drawn in regard to this altitude range. The influence of baseline curvature diminishes rapidly above about 20 km and does not affect our conclusion that the amount of stratospheric N_2O present during the Antarctic spring is greatly reduced compared to the Antarctic summer or to model predictions. We have not removed the residual asymmetry from the data before analysis. (In our work on ClO we have been able to remove any instrumental curvature by subtracting nighttime data^{13,14}, where the signal intensity has been reduced by the diurnal variation of ClO. However, no diurnal variation is expected in the N_2O signal and we are unable to use that technique here.)

Our observations at McMurdo began during the night of 12–13 September 1986, and continued thereafter at one to six day intervals until 28 October. Figure 1 shows the average of all data, except that from 23 and 24 September, taken during this period; the signal has a peak intensity of only 14% of the N_2O signal measured in Hawaii with the same instrument four months earlier¹⁵. Figure 2 shows spectra from a number of days spread through the run. All of these spectra show a similar broad curvature; the data for 23–24 September show a narrow N_2O line feature. In general, the broad curve suggests the presence of some N_2O at altitudes ~ 20 km and below. The narrow peak observed on 23 and 24 September is produced by a thin layer of N_2O at around 35 km, which is discussed in more detail later in this paper.

Figure 1 contains, for comparison with the data, spectra calculated from N_2O altitude distributions predicted by the model of Jones and Pyle⁵, and by that of Solomon and Garcia (ref. 8 and personal communication). These are intended to encompass the range of predictions from recent models for Antarctica in the season when our data was taken. Figure 1 also contains a spectrum calculated for an altitude distribution based on the available measurements made during the Antarctic summer^{3,9}. (The corresponding altitude distributions are displayed in Fig. 3.) The peak intensities of all these synthetic spectra are larger than we observe by a factor of 2.5 or more;

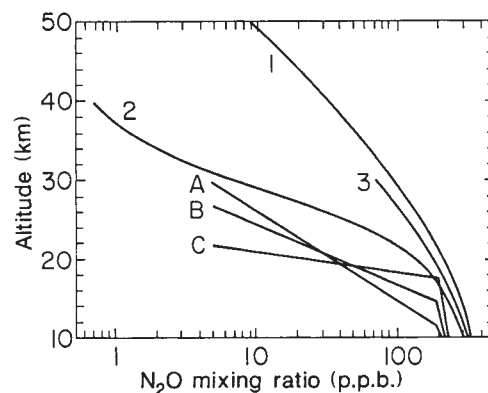


Fig. 3 Curves 1 and 2 are N_2O altitude distributions for 78 S in September predicted by the models of Harwood and Pyle (see ref. 5) and Solomon and Garcia⁸. Other contemporary model predictions that we are aware of lie between curves 1 and 2. Curve 3 displays the average of the Antarctic summertime measurements of Goldan *et al.*³ and Schmeltekopf *et al.*⁹ for altitudes up to 30 km. The synthetic spectra labelled 1, 2, and 3 in Fig. 1 correspond to these altitude distributions. The three curves designated by letter are examples of distributions that produce spectra that are in reasonable agreement with the broad curvature of the data displayed in Figs 1 and 2. (These curves specifically do not include the thin layer of N_2O at about 35 km altitude which is responsible for the peaks seen in the spectra for 23 and 24 September in Fig. 2.) These curves are based on the assumptions discussed in the text. The altitude distribution cannot decrease less rapidly than shown in curve A without producing a spectrum that is noticeably more peaked than the data. Curve C shows that if the mixing ratio is maintained as large as 200 p.p.b.v. up to 18 km, the distribution must decrease very rapidly at higher altitudes.

to obtain reasonable agreement with our data, we would have to divide the altitude distribution from the model of Solomon and Garcia⁸, which predicts the lowest mixing ratios of any model that we are aware of, by a factor of 3 at all altitudes above 15 km. Alternatively, if the altitude distribution of Solomon and Garcia is adjusted downward by 4 km, it will produce a spectrum that agrees reasonably with the data.

Figure 2 shows no secular increase in the observed signal intensity between 12 September and 28 October. This appears to contradict the work of Tung *et al.*^{16,17}, which predicts an upwelling of, and therefore an increase in, the stratospheric abundance of trace species such as N_2O during the very time that these observations were made.

We have developed a simple parameterization in order to determine the major characteristics of the Antarctic N_2O altitude distribution in September, given the limited spectral information available. We have assumed, based on the introductory discussion and references, that the mixing ratio is 300 p.p.b.v. at the surface and decreases slowly to 200 p.p.b.v. at some altitude Z_0 . We have further assumed that, above Z_0 , the mixing ratio decreases exponentially with some scale height Z_s . This second assumption is based on the observation that the N_2O mixing ratio is generally found to decrease in a roughly exponential fashion in the middle stratosphere. We have found distributions of this nature that produce spectra compatible with our data. Three such distributions are shown in Fig. 3, along with two distributions that indicate the range of predictions from current models, and the average of the distributions measured in the Antarctic summer. Curve A decreases with a scale height of 4.7 km above $Z_0 = 12$ km, and produces a spectrum with a hint of a central peak. If the scale height is significantly larger than 4.7 km, the computed line-shape becomes increasingly triangular and qualitatively incompatible with the data shown in Figs 1 and 2. In curve C, the mixing ratio decreases so rapidly above $Z_0 = 18$ km that there is no significant contribution to the spec-

trum from that region. However, if Z_0 is increased further, to 19 km, the corresponding spectrum is quantitatively incompatible with the data; its intensity becomes too large. This curve shows that mixing ratios similar to those measured in summer can only extend to altitudes around 18 km if the mixing ratio decreases unrealistically rapidly at higher altitudes. Curve B is an example of an intermediate case. The three curves cross at 20 km; at this altitude they have a mixing ratio which is a factor of 5 less than the summertime measurements^{3,9}, a factor of 2 less than predicted by the Solomon and Garcia model, and a factor of 6 less than the model results reported. Curve A, which has the largest mixing ratios of the three above 25 km, is less than the summer measurements and the Jones and Pyle model by a factor of 10 at 25 km, and less than the Solomon and Garcia model by a factor of 3. Curves A–C are upper limits in the sense that we cannot rule out lower mixing ratios because of possible weak instrumental curvature.

We have separately analysed the average of the data measured during the event of 23 and 24 September, and find that the observed shape is consistent with N_2O emission from a thin layer with a peak mixing ratio of ~ 40 p.p.b.v. centred at an altitude of around 35 km, in addition to the low altitude N_2O discussed above. This mixing ratio is comparable to the value found at 35 km at lower latitudes. There is no evidence in these data for amounts of N_2O in the 20–30 km range that are greater than those seen in the other observations in this series. Our mm-wave ozone observations¹⁸ during this same event show that a simultaneous increase in ozone took place in the 30–40 km range. Satellite maps show that temperature increased throughout the range 15–40 km during 23 and 24 September over McMurdo. The total column of nitrogen dioxide also increased sharply on 23 September¹⁹ but did not decrease on 24 September. An analysis of potential vorticity by the British Meteorological Office (R. L. Jones, personal communication) shows that, in the 35–40 km altitude range, an intrusion of air which originated at 50° S occurred on these dates which presumably carried mid-latitude amounts of N_2O and ozone with it.

The extent of the region over which the low-altitude N_2O is depleted is indicated by the fact that we find only small amounts of it in 24 observations spread over 2 months. These observations have effectively sampled a substantial number of points spread over the ozone depletion region, which was constantly moving with respect to our site at McMurdo²⁰. N_2O was low even when McMurdo was on the fringes of the region where the ozone was depleted down to 250 Dobson units (DU) or less; thus the geographical extent of the N_2O depletion region appears to be as large as, or perhaps larger than, the ozone hole.

There is no evidence from any models of which we are aware to suggest a precipitous decline in N_2O either induced by transport or by chemistry, during the Antarctic winter. Indeed, as N_2O is destroyed almost solely by photolysis and by reaction with O(1D) in the middle and upper stratosphere, one would expect a low N_2O loss rate during the winter months of darkness.

The particular dynamical aspects of the Antarctic vortex are rather simply treated in the models cited in this paper, because their emphasis has not been on modelling the southern vortex behaviour but rather on developing generally reliable estimates of transport coefficients on a global basis. Thus, the failure of these models to predict extremely low springtime N_2O in Antarctica is not surprising if the cause involves dynamics peculiar to the Antarctic atmosphere. If so, the global models may be seriously flawed when specifically applied to chemical explanations of the Antarctic ozone hole, because a local transport perturbation which affects the mixing ratio of one species must almost certainly also affect those of other species.

Using a three-dimensional model with emphasis on dynamical treatment, Mahlman and Fels²¹ have explicitly sought to investigate the effects of reduced planetary-scale disturbances on Antarctic ozone. The model predicts ozone depletion in late winter and early spring but at the same time predicts a decrease

in the latitudinal gradient in N_2O , which would produce an increase in the Antarctic N_2O concentration.

In our papers discussing our measurements of ClO in Antarctica^{13,14}, we present data showing chlorine to be strongly involved in the ozone depletion mechanism at low altitudes. The very low mixing ratios of N_2O that we report here further demonstrate the unusual and unexpected nature of the Antarctic springtime stratosphere.

We thank the NSF's Division of Polar Programs for its support through all phases of our Antarctic work, and the many individuals in Antarctic Services, Inc., the US Navy Support Force and the NSF whose interest and enthusiasm smoothed our way. We thank Dr R. L. Jones for supplying us with information on the trajectory of air over McMurdo during 23–24 September 1986, and NASA's Upper Atmospheric Research Program and the Chemical Manufacturer's Association for the support which made it possible to carry out this research.

Received 19 May; accepted 17 December 1987.

1. World Meteorological Organization/NASA, *The Stratosphere, 1981; Theory and Measurement* (Report No. 16 of the WMO Global Ozone Research and Monitoring Project, NASA Headquarters, Washington, 1986).
2. Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207–210 (1985).
3. Goldan, P. D., Kuster, W. C., Albritton, D. L., & Schmeltkeopf, A. L. *J. geophys. Res.* **85**, 413–423 (1980) **86**, 5385–5386 (correction) (1981).
4. Vedder, J. F., Inn, E. C. Y., Tyson, B. J., Boitnott, C. A. & O'Hara, D. *J. geophys. Res.* **86**, 7363–7368 (1981).
5. Jones, R. L. & Pyle, J. A. *J. geophys. Res.* **89**, 5263–5279 (1984).
6. Mahlman, J. D., Levy, H. II & Moxim, W. J. *J. geophys. Res.* **91**, 2681–2686 (1986).
7. Ko, M. K. W., Tung, K. K., Weinstein, D. K. & Sze, N. D. *J. geophys. Res.* **90**, 2313–2329 (1985).
8. Solomon, S. & Garcia, R. R. *J. geophys. Res.* **89**, 11633–11644 (1984).
9. Schmeltkeopf, A. L. *et al. J. Atmos. Sci.* **34**, 729–736 (1977).
10. Goldman, A., Fernald, F. G., Murcray, F. J., Murcray, F. H. & Murcray, D. G. *J. quant. Spect. Rad. Trans.* **29**, 189–204 (1983).
11. Parrish, A., deZafra, R. L., Solomon, P. M. & Barrett, J. W. *Radio Sci.* (in the press).
12. Hofmann, D. J., Harder, J. W., Rolf, S. R. & Rosen, J. M. *Nature* **326**, 59 (1987).
13. deZafra, R. L. *et al. Nature* **328**, 408–411 (1987).
14. Solomon, P. M. *et al. Nature* **328**, 411–414 (1987).
15. Connor, B. J. *et al. Geophys. Res. Lett.* **14**, 1254 (1987).
16. Tung, K. K., Ko, M. K. W., Rodriguez, J. M. & Sze, N. D. *Nature* **322**, 811 (1986).
17. Tung, K. K. *Geophys. Res. Lett.* **13**, 1308–1311 (1986).
18. Connor, B. *et al. J. geophys. Res.* **92**, 13221–13230 (1987).
19. Mount, G., Sanders, R., Schmeltkeopf, A. & Solomon, S. *J. geophys. Res.* **92**, 8320–8327 (1987).
20. Krueger, A. J., Schoeberl, M. R. & Stolarski, R. S. *Geophys. Res. Lett.* **14**, 527–530 (1987).
21. Mahlman, J. D. & Fels, S. B. *Geophys. Res. Lett.* **13**, 1316–1319 (1986).

Superconductivity in the rare-earth-free Tl–Ba–Cu–O system above liquid-nitrogen temperature

Z. Z. Sheng & A. M. Hermann

Department of Physics, University of Arkansas, Fayetteville, Arkansas 72701, USA

The initial discovery by Bednorz and Muller¹ of 35-K superconductivity in the La–Ba–Cu–O system has stimulated worldwide activity in searching for higher-temperature superconductors. Elemental substitution has proved to be most effective in raising transition temperature. Substitution of Sr for Ba has produced 40-K superconductivity^{2–5}, and substitution of Y for La has produced a new high-temperature superconductor with transition temperature above liquid-nitrogen temperature⁶. A class of superconducting compounds of the form $RBa_2Cu_3O_{7-x}$ has been explored by further substitutions of other rare earths (Y is considered in the rare-earth [R] category here) for Y^{7–13}. To date, a rare earth, an alkaline earth, copper and oxygen have been required for all high-temperature superconductors^{14,15}. (Zhang *et al.*¹⁴ reported 90-K superconductivity in the Th–Ba–Pb(Zr)–Cu–O system. Pan *et al.*¹⁵ reported 50-K superconductivity in the Y–Ba–Ag–O system. As Th is a member of the actinide series which belongs to the same Group 3B in the periodic table as the lanthanide series and Ag

belongs to the same Group 1B as Cu, high-temperature superconductors are still thought to be closed in the Group 3B–Group 2A–Group 1B–oxygen system.) Only partial substitutions have led to superconductors, but with no significant rise of transition temperature (the only exception is 40-K superconductivity in $\text{La}_2\text{CuO}_{4-x}$, refs 16, 17). Here we report superconductivity in the rare-earth-free Tl-Ba-Cu-O system. We have observed sharp drops of resistance starting above 90 K with zero resistance at 81 K in this system. Magnetic measurements have confirmed that these sharp drops of resistance in the Tl-Ba-Cu-O samples originate from superconductivity. The samples are stable in air for at least two months, and their preparation is easily reproduced.

Historically, when 90-K superconductivity in the Y-Ba-Cu-O system was announced (and the superconducting compound $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ was identified¹⁸), it became clear immediately that substitutions of other rare earths for Y would produce superconductors. Two approaches for rare-earth substitution were taken. One approach was the choice of rare earths with small or no magnetic moment such as Lu and Yb, as a small concentration of magnetic ions is known to depress significantly or even to destroy superconductivity in conventional superconductors. Another approach was the choice of rare earths with ionic radius close to that of Y, such as Gd and Ho, in spite of their large magnetic moment. After all the rare earths were substituted^{7–13}, it could be concluded that: (1) substitution of magnetic rare earths for Y in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ did not influence significantly the transition temperature; (2) ionic radius of rare earths was important to superconductive behaviour of the $\text{RBa}_2\text{Cu}_3\text{O}_{7-x}$ compounds; for example, substitution of La (the rare earth with largest ionic radius) for Y produced superconductivity with a relatively low transition temperature, usually below liquid-nitrogen temperature, and the Lu-containing superconductor (Lu is the rare earth with smallest ionic radius) has not been prepared in single phase so far because the $\text{LuBa}_2\text{Cu}_3\text{O}_{7-x}$ phase is difficult to form; and (3) the valence of the substituted rare earth was crucial to the superconductivity of the $\text{RBa}_2\text{Cu}_3\text{O}_{7-x}$ compounds (Ce, Pr and Tb with both 3+ and 4+ valences did not form superconducting compounds). These facts strongly suggest that the valence and radius of rare-earth ions are more vital to the formation of rare-earth-containing superconductors than are local magnetic effects.

Logically, one would expect that other elements with 3+ valence and similar ionic radius could also form similar superconducting compounds. By inspecting the periodic table, the elements in Group 3A are also characterized by a 3+ valence. In particular, thallium, the heaviest element in Group 3A, has an ionic radius in 3+ valence state of 0.95 Å, similar to those of rare earths (1.016–0.85 Å) and the same as that of Eu^{3+} (ref. 19). Thallium would thus appear to be the best candidate which could also form a similar superconducting compound. It was these considerations that motivated us to completely replace rare earths in the R-Ba-Cu-O system with thallium.

Samples were prepared using 99.999% pure Tl_2O_3 , 99.999% pure BaCO_3 and certificate reagent CuO . Because Tl_2O_3 has a relatively low melting point of 717 °C and starts to decompose at a temperature as low as 100 °C^{19,20}, typical procedures used for preparing $\text{RBa}_2\text{Cu}_3\text{O}_{7-x}$ superconductors do not apply to the Tl-Ba-Cu-O system. In our preparation of the Tl-Ba-Cu-O superconductor, we used short high-temperature heating and quenching. High temperature is required for the initiation of the formation reaction of the superconducting compound; short-duration heating minimizes the decomposition and evaporation of Tl_2O_3 ; quenching stabilizes the superconducting phase. Because melt-solid reactions take place much more readily than conventional solid-solid reactions, in some procedures, we also used principles of the melt process (based on melt-solid reactions) that we developed for preparing high-quality melt-processible R-Ba-Cu-O superconductors^{21,22}. The following is a typical procedure for preparing Tl-Ba-Cu-O samples. Appropriate amounts of BaCO_3 and CuO were mixed and

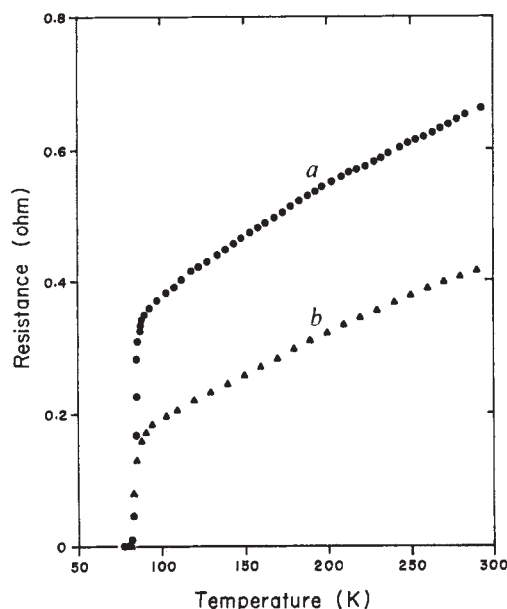


Fig. 1 Variations of resistance with temperature for samples *a*, $\text{Tl}_2\text{Ba}_2\text{Cu}_3\text{O}_{8+x}$ and *b*, $\text{TlBaCu}_3\text{O}_{5.5+x}$.

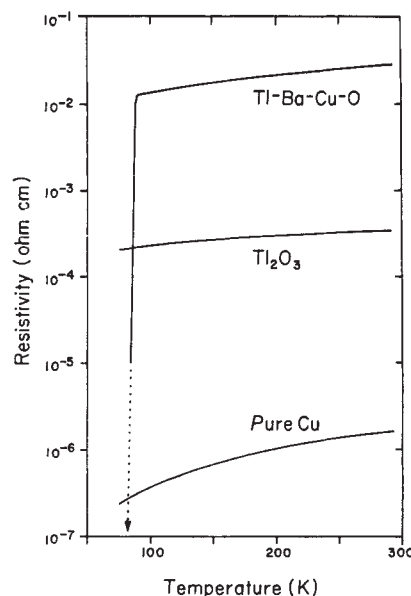


Fig. 2 Variation of resistivity with temperature for a $\text{Tl}_2\text{Ba}_2\text{Cu}_3\text{O}_{8+x}$ sample.

ground with an agate mortar and pestle, and heated in air at 925 °C for >24 h with several intermediate grindings to obtain a uniform black Ba–Cu oxide powder (for example, BaCu_3O_4 or $\text{Ba}_2\text{Cu}_3\text{O}_5$). The resulting Ba–Cu oxide powder was mixed with an appropriate amount of Tl_2O_3 (depending on the desired stoichiometry), completely ground, and pressed into a pellet with a diameter of 7 mm and a thickness of 1–2 mm. The pellet was then put into a tube furnace which had been heated to 880–910 °C, and was heated for 2–5 min in flowing oxygen. As soon as it had slightly melted, the sample was taken from the furnace and quenched in air to room temperature. We noted by visual inspection that Tl_2O_3 had partially volatilized as black smoke, part had become a light yellow liquid, and part had reacted with Ba–Cu oxide forming a black (partially melted) porous material. Because the superconducting $\text{RBa}_2\text{Cu}_3\text{O}_{7-x}$ compounds have special stability in the presence of oxygen, either in structure or in energy, and small amounts of rare earth

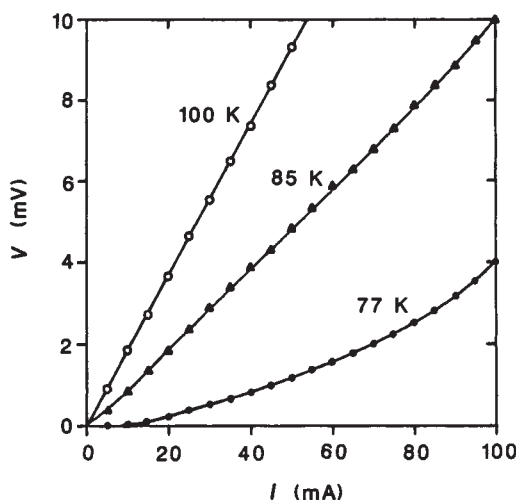


Fig. 3 Current-voltage curves at 77, 85 and 100 K for a $\text{Tl}_2\text{Ba}_2\text{Cu}_3\text{O}_{8+x}$ sample.

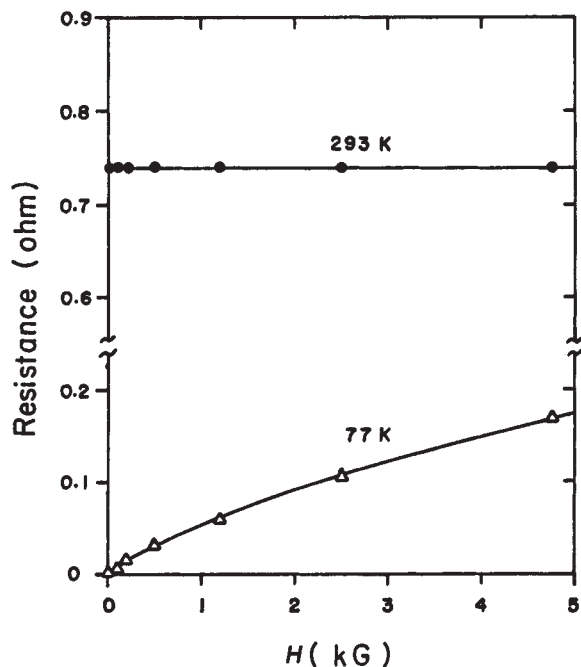


Fig. 4 Variations of resistance with magnetic field at 77 K and room temperature for a $\text{Tl}_2\text{Ba}_2\text{Cu}_3\text{O}_{8+x}$ sample.

mixed with Ba-Cu oxides will lead to formation of these compounds²¹, special attention was paid to avoid contamination of rare-earth elements in all experiments.

Resistance was measured by the standard four-probe technique with silver paint used for the contacts. The samples were of typical dimensions $7 \times 3 \times 1$ mm. All measurements were carried out in a simple liquid-nitrogen Dewar. The variation of resistance with temperature of Tl-Ba-Cu-O samples is similar to that observed for typical $\text{RBa}_2\text{Cu}_3\text{O}_{7-x}$ samples: resistance decreases nearly linearly with decreasing temperature above the onset temperature, and then sharply drops, and the ratio of resistance at room temperature to that at onset temperature is ~ 2 . Figure 1 shows variations of resistance for samples *a* and *b*, which have nominal compositions of $\text{Tl}_2\text{Ba}_2\text{Cu}_3\text{O}_{8+x}$ and $\text{TlBaCu}_3\text{O}_{5.5+x}$, respectively. Both samples have onset temperatures of >90 K and midpoints of ~ 85 K, and reach zero resistance at 81 K.

Figure 2 shows the variation of resistivity with temperature

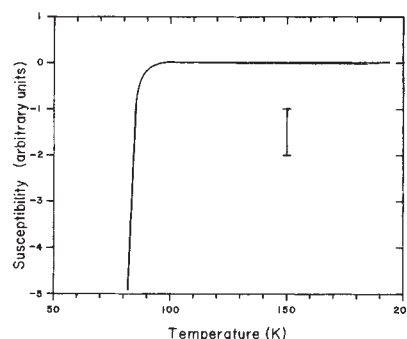


Fig. 5 Low frequency a.c. susceptibility of a sample of $\text{Tl}_{1.2}\text{BaCu}_3\text{O}_{5.6+x}$. The data from four runs were averaged to give the solid curve (the bar denotes the spread). The zero was chosen arbitrarily.

for a $\text{Tl}_2\text{Ba}_2\text{Cu}_3\text{O}_{8+x}$ sample. The porosity and multi-phase nature of the sample limit accurate measurements of resistivity. The actual resistivity of the current-carrying phase will certainly be much less than the calculated values. For comparison, resistivity-temperature plots for pure copper^{19,23} and a sintered Tl_2O_3 sample are also shown in Fig. 2. The average resistivity of the multi-phase sample $\text{Tl}_2\text{Ba}_2\text{Cu}_3\text{O}_{8+x}$ at 77 K has decreased to $<10^{-8} \Omega \text{ cm}$, which is much lower than that of pure copper at the same temperature. This provides strong evidence for the existence of a high-temperature superconducting phase in the sample²⁴. It is interesting to note that the Tl_2O_3 has low room-temperature resistance and is metallic above liquid-nitrogen temperatures.

Figure 3 shows I - V curves for a sample of $\text{Tl}_2\text{Ba}_2\text{Cu}_3\text{O}_{8+x}$ at 77 K, 85 K and 100 K. This sample is normal at 100 K, and thus the I - V curve is linear, as seen for normal materials. At 85 K the I - V curve is characteristic of a mixed-state superconductor. At 77 K, at which temperature the sample was in a superconducting state, the resistance of this sample is zero up to a critical current density of $\sim 200 \text{ mA cm}^{-2}$. These I - V curves are consistent with the expectation that segments in a superconducting path connected serially to a normal path turn into normal states one-by-one under the increasing current. This leads to an increase in the overall resistance²⁵.

Figure 4 shows variations of resistance of a sample of $\text{Tl}_2\text{Ba}_2\text{Cu}_3\text{O}_{8+x}$ with magnetic field at 77 K and room temperature. It can be seen that resistance increases with transverse magnetic field in the superconducting state, and does not change in the normal state. These results clearly show that the loss of resistance of the Tl-Ba-Cu-O samples arises from the presence of a superconducting phase²⁶.

Qualitative magnetic examinations on Tl-Ba-Cu-O samples, which have been proved to be very effective for quick checks of superconductivity, showed an onset of a diamagnetic component at ~ 90 K. Figure 5 shows the low-frequency a.c. susceptibility for a sample $\text{Tl}_{1.2}\text{BaCu}_3\text{O}_{5.6+x}$ as measured by the technique of Norton²⁷. The transition appears incomplete at 82 K, but the precipitous drop near the zero-resistance value (~ 81 K) is clear. The signal size is an order of magnitude smaller than that for a typical $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ sample, as would be expected for a bulk mixed-phase superconductor with about 5% volume of superconducting phase.

Preliminary results of X-ray diffraction analysis showed that, compared to the starting materials Tl_2O_3 and Ba-Cu oxide, the superconducting samples have two additional diffraction peaks, one at 2θ of just below 33° (characteristic of an $\text{RBa}_2\text{Cu}_3\text{O}_{7-x}$ perovskite), and one at 57° , which are believed to be characteristic of the phase(s) responsible for the superconductivity. Detailed structure analyses are in progress and the results will be published elsewhere.

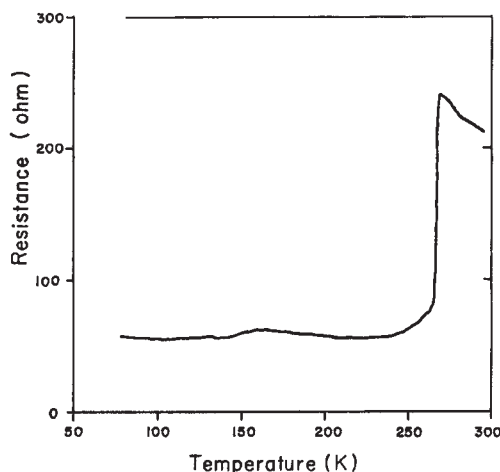


Fig. 6 Resistance-temperature variation for a Tl-Ba-Cu-O sample showing a sharp drop of the resistance at 267 K.

Finally, we would like to note that we have observed a sharp drop of resistance at 267 K for one sample of this new superconductive Tl-Ba-Cu-O system as shown in Fig. 5. The results were reproduced during four temperature cycles between room temperature and liquid-nitrogen temperature. There are many factors that can cause sharp drops of resistance, especially for multi-phase samples with high resistance. However, for a strong electron-phonon interaction as required for example by BCS theory, high resistance in the normal state is expected (Fig. 2 provides an example of this). Preliminary experiments on the Tl-Sr-Cu-O system have also shown interesting results, including further suggestions of a higher-temperature superconductive phase. The variation of resistance with temperature for the samples in this system is rather similar to that seen in Fig. 2 of Bednorz and Muller's initial paper¹ announcing high-temperature superconductivity in the La-Ba-Cu-O system. However, the drops of resistance in the La-Ba-Cu-O system appear at ~30 K, whereas in the Tl-Sr-Cu-O system they appear between 110 and 200 K.

We thank Mark J. Dreiling of the Phillips Petroleum Company for providing the low-frequency a.c. susceptibility data.

The effect of density on critical current and oxygen stoichiometry of $\text{YBa}_2\text{Cu}_3\text{O}_x$ superconductors

N. McN. Alford*, W. J. Clegg*, M. A. Harmer*, J. D. Birchall*, K. Kendall* & D. H. Jones†

*ICI Advanced Materials, Runcorn, Cheshire WA7 4QD, UK

†Physics Department, Liverpool University, Oxford Street, Liverpool L69 3BX, UK

It has been observed that "the most important parameter of a superconductor... is the critical current density J_c " (ref. 1). In thin films or single crystals of $\text{YBa}_2\text{Cu}_3\text{O}_x$, J_c is high at $\sim 10^6 \text{ A cm}^{-2}$ but in bulk materials J_c is much lower, typically 10^2 – 10^3 A cm^{-2} (ref. 4). The sources of variation in J_c are important but vary. Chemical contamination such as SiO_2 impurities is known to reduce J_c ⁵ but even in the absence of such contaminants J_c varies widely. Some groups have experienced difficulty in maintaining high ($1,000 \text{ A cm}^{-2}$ in bulk ceramic) current densities and sometimes even zero resistance at 77 K at high densities of the ceramic ($>6.2 \text{ g cm}^{-3}$). J_c is known to be anisotropic being greater in the a - b planes and least in the c -axis direction. In thin films a thirty-fold difference in J_c was reported demonstrating the anisotropy⁶. In bulk materials J_c has been increased to $7.5 \times 10^3 \text{ A cm}^{-2}$ by melt textured growth of the ceramic in order to induce preferred orientation⁷. Here we report the factors controlling J_c in sintered bulk ceramics that display no preferred orientation. We examine the relation between sintered density, oxygen content and J_c and show how J_c is strongly dependent on both the volume fraction of solid and on the oxygen stoichiometry, both being strongly influenced by the ceramic density. The problem we address is how to optimize the solid connectivity and simultaneously maintain a high oxygen content thereby maximizing J_c in anisotropic bulk ceramics.

$\text{YBa}_2\text{Cu}_3\text{O}_x$ was prepared by calcining BaCO_3 (99.9%), Y_2O_3 (99.9%) and CuO (99.9%) at 900 °C for 10 h followed by comminution in a polymer vibro energy mill using ZrO_2 media. The resultant powder had a surface area of $3 \text{ m}^2 \text{ g}^{-1}$ yielding an equivalent spherical diameter of 0.3 μm . This powder was processed by viscous processing techniques⁸ using non-aqueous polymers and solvents to avoid deleterious reaction of the $\text{YBa}_2\text{Cu}_3\text{O}_x$. The mixture was extruded in a ram extruder to form rods 1 mm in diameter. Sintering took place in a silica tube furnace under flowing oxygen. The sintering temperature was varied between 900–940 °C in order to vary the sample density. Furnace cooling varied from 1 °C min^{-1} to room tem-

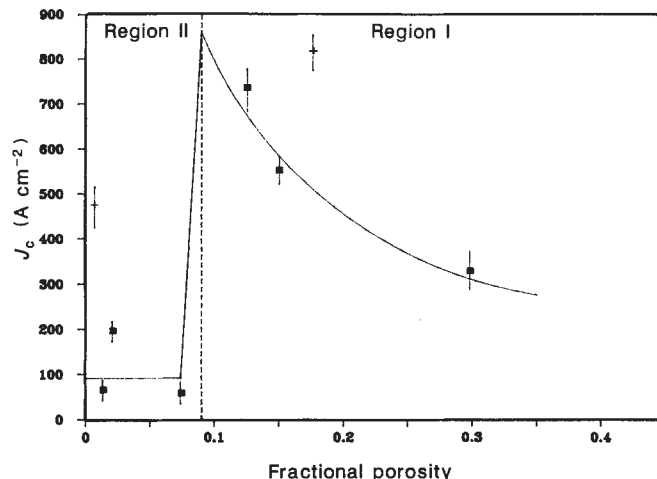


Fig. 1 J_c as it varies with fractional porosity. ■, Samples cooled at 1 °C min^{-1} ; +, samples annealed at 400 °C in O_2 . Error bars are one standard deviation.

Received 20 January; accepted 10 February 1988.

- Bednorz, J. G. & Muller, K. A. *Z. Phys.* **B64**, 189–193 (1986).
- Chu, C. W. *et al. Phys. Rev. Lett.* **58**, 405–407 (1987).
- Chu, C. W. *et al. Science* **235**, 567–569 (1987).
- Cava, J. *et al. Phys. Rev. Lett.* **58**, 408–410 (1987).
- Tarascon, J. M. *et al. Science* **235**, 1373–1376 (1987).
- Wu, M. K. *et al. Phys. Rev. Lett.* **58**, 908–910 (1987).
- Murphy, D. *et al. Phys. Rev. Lett.* **58**, 1888–1890 (1987).
- Hor, P. H. *et al. Phys. Rev. Lett.* **58**, 1891–1894 (1987).
- Moodenbaugh, M. *et al. Phys. Rev. Lett.* **58**, 1885–1887 (1987).
- Hosoya, S. *et al. Jap. J. appl. Phys.* **26**, L325–L326 (1987).
- Kitazawa, K. *et al. Jap. J. appl. Phys.* **26**, L339–L341 (1987).
- Hikami, S. *et al. Jap. J. appl. Phys.* **26**, L347–L348 (1987).
- Hosoya, S. *et al. Jap. J. appl. Phys.* **26**, L456–L457 (1987).
- Zhang, Y. L. *et al. preprint* (1987).
- Pan, K. K. *et al. Phys. Lett. A* **125**, 147–148 (1987).
- Grant, P. M. *et al. Phys. Rev. Lett.* **58**, 2482–2485 (1987).
- Michel, C. *et al. Z. Phys.* **B68**, 421–423 (1987).
- Cava, J. *et al. Phys. Rev. Lett.* **58**, 1676–1679 (1987).
- Weast, R. C. W. & Astle, M. J. (eds) in *Handbook of Physics and Chemistry*, 62nd edn, B-156, F-175 (CRC, Boca Raton, 1981).
- Bailar, J. C. Jr *et al. Comprehensive Inorganic Chemistry* (Pergamon, Oxford, 1973).
- Hermann, A. M. & Sheng, Z. Z. *Appl. Phys. Lett.* **51**, 1854–1856 (1987).
- Hermann, R. *et al. Appl. Phys. Commun.* **7**, 275–283 (1987).
- White, G. K. & Woods, S. B. *Phil. Trans. R. Soc. A* **251**, 272 (1959).
- Ovshinsky, S. R. *et al. Phys. Rev. Lett.* **58**, 2579–2581 (1987).
- Kagoshima, S. *et al. Jap. J. appl. Phys.* **26**, L318–L319 (1987).
- Matsushita, A. *et al. Jap. J. appl. Phys.* **26**, L332–L333 (1987).
- Norton, M. L. *J. Phys. E. Sci. Instrum.* **19**, 268–270 (1986).

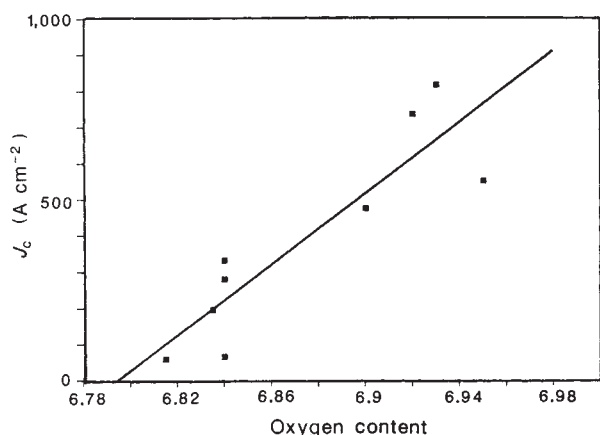


Fig. 2 J_c as it varies with oxygen content.

perature, and 1°C min^{-1} with anneals at 400°C for 15–30 h. Each sample batch contained 10–15 rods of ~ 1 mm diameter by 50 mm in length. Oxygen content was measured by thermogravimetric analysis (Stanton Redcroft). The critical current measurements were made using a Cryogenic Consultant TS120A. Pressure contacts with a resistance of $<1\ \Omega$ were made to rods of ~ 0.8 mm diameter. The apparatus was immersed in liquid nitrogen, which was found to disperse adequately the maximum heat produced at the current contacts. A total of 129 rods were analysed for density, critical current and oxygen content.

Figure 1 shows the effect of sample density on J_c and may be separated into two regions. In region I the oxygen content of the samples is almost identical with $x = 6.937 \pm 0.015$ and J_c increases due to the increased connectivity of the ceramic structure as density increases from $\rho = 0.7\ \rho_{\text{th}}$ (theoretical density) to $\rho = 0.87\ \rho_{\text{th}}$. The values from region I in Fig. 1 may be used in a simple rule of mixtures approach to estimate J_c^{max} at zero porosity

$$J_c = J_c^{\text{max}}(1 - P)^3 \quad (1)$$

where J_c^{max} is the critical current at zero pore volume, P is fractional pore volume and J_c is the measured value of J_c . Equation (1) yields a value of approximately $1,000\ \text{A cm}^{-2}$ as an estimate for fully dense material.

From Fig. 1 it can be seen that for densities above $0.92\ \rho_{\text{th}}$ (that is, in region II) the critical current falls suddenly to about 1/5 of the value at $0.85\ \rho_{\text{th}}$.

This is associated with the material becoming impermeable as the pores close off and with a decrease in the oxygen value determined by thermogravimetric analysis from 6.95 to 6.82. This is consistent with the view that once pore closure has occurred oxygen must diffuse through the bulk during the cool to room temperature and that this bulk diffusion is relatively slow. An impermeable superconducting sample is therefore made up of an oxygen-rich surface with an oxygen-poor core.

If an increasing current is passed along such a rod below the transition temperature, T_c , then initially the current will flow in both the core and the sheath. As the current rises the initial current density of the core will be exceeded and current will flow only in the sheath. We have assumed that there are no thermal effects due to heating in the core once this is above its critical current density because the current will flow only in the sheath.

On this basis, we assume that to a first approximation the critical current of the impermeable rod, J_c^{I} , is related to the area fraction of the sheath, A_s , and the critical current of a fully dense fully oxygenated material, J_c^{max} :

$$J_c^{\text{I}} = A_s J_c^{\text{max}} \quad (2)$$

J_c^{max} must be close to the value of the $0.85\ \rho_{\text{th}}$ dense material.

If the critical current falls by a factor of 5, that is $J_c^{\text{max}}/J_c^{\text{I}} = 5$, then we can calculate the diameter of the sheath. For our 0.8-mm rods this gives an approximate thickness for the sheath of only $40\ \mu\text{m}$. Despite this the rods are still superconducting and show a resistance falling to zero at $\sim 90\ \text{K}$.

To increase the depth of the oxygen-rich surface, samples were cooled under oxygen at 1°C min^{-1} to 400°C and then held at 400°C for 15 hours before cooling to room temperature. The oxygen level, x , rose to 6.93, and J_c to $475 \pm 71\ \text{A cm}^{-2}$.

Further evidence for this effect is shown in Fig. 2 where the critical current density is plotted against the oxygen concentration. We can see how this arises if we assume that the measured oxygen concentration is given by

$$x = A_c x_c + A_s x_s \quad (3)$$

where A_c , A_s are the areal fraction of the core and sheath respectively and x_c , x_s are the respective oxygen concentrations.

$$A_c = (1 - A_s) \quad (4)$$

then combining equations (2), (3) and (4) gives

$$J_c^{\text{I}} = \frac{x J_c^{\text{max}}}{x_s - x_c} - \frac{x_c J_c^{\text{max}}}{x_s - x_c} \quad (5)$$

From equation (5) it can be seen that the initial current density of the impermeable rod J_c^{I} is a linear function of the oxygen concentration x as J_c^{max} , x_s and x_c are all approximately constant.

It is worth noting that if we extrapolate the curve to $x = 7$ then J_c would give a value of $\sim 1,000\ \text{A cm}^{-2}$ in agreement with the prediction from equation (1). Prospects for increasing J_c in the bulk ceramic will require techniques such as grain alignment during processing and firing.

We thank G. Dolman, J. Whitehouse and J. Chadwick for useful discussions and assistance with the experimental work.

Received 5 January; accepted 29 January 1988.

1. Dew-Hughes, D. *Phil Mag B* **55**, 459–479 (1987).
2. Ekin, J. W. *Adv. Ceram. Mat.* **2**, 586–591 (1987).
3. Alford, N. McN. *et al. J. Mater. Sci.* (in the press).
4. Glowacki, B. A. & Evetts, J. E. *Proc. Mater. Res. Soc. 1987 Fall Mtg.* (in the press).
5. Newcomb, S. B., Glowacki, B. A., Evetts, J. E. & Stobbs, W. M. *Proc. 14th Br. Ceram. Soc.* (in the press).
6. *High Tech Mater. Alert* Vol. 4(6), 3 (1987).
7. *Superconductor Week* 1(19), (1987).
8. Alford, N. McN., Birchall, J. D., Howard, A. J., Kendall, K. & Raistrick, J. H. *European Patent Publication* No. 183453 (1986).

New and unexpected reactivity of saturated fluorocarbons

David D. MacNicol & Colin D. Robertson

Department of Chemistry, University of Glasgow,
Glasgow G12 8QQ, UK

The carbon–fluorine bond, noted for its strength¹, has been of chemical interest² even before Moissan's discovery of elemental fluorine. The most electronegative element, fluorine is uniquely capable of replacing all the hydrogen atoms in every hydrocarbon system³, yielding materials such as polytetrafluoroethylene renowned for their combination of rare physical properties, outstanding chemical inertness and high thermal stability⁴, with significant modern technological applications^{5,6}. Reaction of saturated fluorocarbons requires extreme conditions¹, for example treatment with metals at $\sim 500^\circ\text{C}$ ⁷ or high-energy irradiation⁸. Following an observation by MacNicol and McGregor (unpublished data) of an unprecedented reactivity under mild conditions we now report specifically the efficient transformation of perfluorodecalin (1) to

known host molecules by arenethiolate nucleophiles in dipolar aprotic solvent. This chemistry is not only of fundamental mechanistic interest, but also creates new synthetic possibilities in various fields. For example, preparing novel hosts for inclusion chemistry⁹, a subject attracting much current attention.

Typical results for the reaction of (1), used as a commercial mixture of *cis* and *trans* isomers, with the sodium salts (2a-c) at ambient and higher temperatures, mostly performed in 1,3-dimethylimidazolidin-2-one (DMEU), are summarized in Table 1. The isolated products, octakis(phenylthio)naphthalene (3a) and octakis(*m*-tolylthio)naphthalene (3b) had a melting point at 205–206 °C and 156–157 °C respectively (both as unsolvated forms)^{10,11}, these materials also had spectroscopic properties (infrared, ¹H NMR, ¹³C NMR, and mass spectrometry) identical with those of authentic materials prepared from octafluoronaphthalene, the structures of which have been unambiguously established by single-crystal X-ray crystallography^{10,11}. Compound (3c), octakis(*m*-methoxyphenylthio)naphthalene had spectroscopic properties (infrared, ¹H NMR, ¹³C NMR and mass spectrometry) fully in accord with its assigned structure, these being identical to those obtained from material prepared by the authors (unpublished data) from octafluoronaphthalene. For the reaction of (2a) with (1) at ambient temperature little reaction is observed as red coloration, after one month. The process is significantly accelerated by increasing the reaction temperature to ~60–70 °C without promoting by-product formation. Above this temperature replacement of arylthio substituents by hydrogen occurs, as indicated by ¹H NMR, ¹³C NMR and mass spectrometry.

We are currently considering the mechanistic aspects of this novel reaction, the crucial first step of which may involve nucleophilic attack on carbon or fluorine, or single electron transfer (SET) to form a radical anion, by the sulphur nucleophile. In this case nucleophilic attack on a carbon atom of (1) appears unlikely¹², and from the literature it appears that either of the other two pathways is more probable. A number of recent studies have shown that haloalkylfluorocarbons¹³, normally considered inert, readily undergo substitution through attack on chlorine and bromine by nucleophiles (for example, sulphur, oxygen and nitrogen-based systems). The halophilic attack has in different cases been assigned to operate by both anionic and SET mechanisms. The work here may extend these findings to fluorine in such systems as (1). The high electron affinity of saturated cyclic fluorocarbons suggests generation of a radical anion could take place by a different mechanism¹⁴. It is interesting to note that the decomposition of *n*-perfluoroalkane in Na/NH₃(l) by an unknown pathway¹⁵ may be related despite the absence of reactivity of this class of molecule under present conditions.

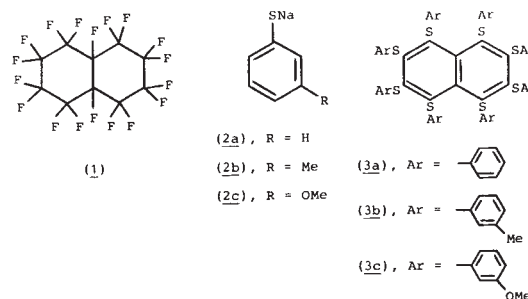


Fig. 1 Structure of the unsaturated fluorocarbon (1), the sodium salts (2a-c) and the isolated products (3a-b) of the reactions between (1) and (2a-c).

In the new reaction nucleophilic attack on a tertiary fluorine (see below) would lead to the bridgehead carbanion C₁₀F₁₇⁻ (4) directly, recently isolated as its tris(dimethylamino)sulphonium salt¹⁶, which is expected to eliminate fluoride to give the symmetric perfluorobicyclic olefin C₁₀F₁₆ (5) with no vinylic fluorine; equally SET would lead to a radical anion of (1), C₁₀F₁₈^{•-}, capable of eliminating fluoride to give a tertiary carbon radical which could then gain an electron to form (4) and hence (5) or alternatively could combine with arenethiolate (2) to generate the thioether derivative¹⁷ after losing an electron, nucleophilic attack on the sulphur of which again leads to (5) via (4)¹⁸.

Consistent with the possible intermediacy of (5), which might be expected to lead rapidly to the observed products, we have subjected perfluorocyclohexene, a model system, to analogous reaction conditions using (2a) in DMEU. This reacted rapidly at ambient temperature to give a 96% yield of hexakis(phenylthio)benzene¹⁹, C₆(SPh)₆, melting point at 185–186 °C, which was isolated as its 1,1,1-trichloroethane clathrate (host-guest ratio 1:2). Indeed, facile multiple substitution initiated by nucleophilic attack on the double bond in (5) has been demonstrated²⁰ and a recent paper²¹ on the multisubstitution with concomitant introduction of further unsaturation of an acyclic perfluoro-olefin, under similar conditions of excess mercaptide, are in keeping with the above proposal. The importance of a tertiary centre in the molecule is indicated by the failure of perfluorocyclohexane and *n*-perfluorohexane to react under equivalent conditions, whereas perfluoro(methylcyclohexane) and perfluoroperhydrophenanthrene do.

In view of these findings it is important to eliminate the possibility of significant levels of impurities starting the process, especially olefinic residues or olefinic precursors, such as hydrogen-containing material. The starting material was rigorously checked by infrared and Raman spectroscopy, and by ¹H, ¹³C and ¹⁹F NMR, and all of these methods gave no evidence for the presence of hydrogen or unsaturation, while independent gas-liquid chromatographic analysis determined the content of (1) as 95.6% (the remainder being saturated fluorocarbons arising from fluorinolysis and rearrangement occurring during synthesis). In addition the reaction was repeated on a medical-grade sample of (1) which is necessarily supplied free of these highly toxic impurities.

Perfluorodecalins are used in conjugation with other perfluoro-emulsifying agents as artificial blood substitutes²², where they have been found non-toxic. But the above findings may suggest perfluorochemicals cannot now be automatically considered indefinitely inert under all mild conditions, including biological systems, where electron donors and nucleophiles are active. Further work in this laboratory has now also shown that the 'inert' fluorocarbon (1) can be reacted with selenium and oxygen-based nucleophiles, extending the scope of this previously unsuspected reactivity; it is expected that future efforts will lead to functionalized fluoropolymers, potential applications of which have already been identified²³.

Table 1 Reactions of perfluorodecalin (1) with sulphur nucleophiles (2)

Reagent*†	Solvent	Temperature	Time	Product	Yield‡ (%)
(2a)	DMEU	ambient	several months	(3a)	17§
(2a)	DMEU	60–70 °C	10 days	(3a)	65
(2a)	DMF#	60–70 °C	10 days	(3a)	55
(2b)	DMEU	ambient	8 weeks	(3b)	20
(2c)	DMEU	ambient	several months	(3c)	5 ¶

* With respect to (1), 36 molar equivalents were employed.

† The salt was preformed and dried before use.

‡ Yields based on (1) and not optimized. A large quantity of diaryl disulphide was formed in each case. Also, fluoride was detected by ion-selective electrode.

§ Unreacted (1) has an unaltered *cis/trans* ratio, by ¹⁹F NMR.

|| Most of (1) recovered unreacted.

¶ Small amounts of related products were also isolated.

Dimethyl formamide.

We thank the SERC for support (to C.D.R.), ISC Chemicals Ltd. (Bristol) for sample analysis, Dr P. L. Coe (University of Birmingham) for providing some samples, Dr A. Cooper (University of Glasgow) for the Raman spectra, and SERC Mass Spectrometry service (Swansea). The fluorocarbons discussed here are the subject of British Patent Application No. 8720773 (4 September 1987).

Received 2 November 1987; accepted 5 January 1988.

1. Tatlow, J. C. *J. fluor. Chem.* **25**, 99-110 (1984).
2. Banks, R. E. & Tatlow, J. C. *J. fluor. Chem.* **35**, 3-4 (1987).
3. Banks, R. E. & Tatlow, J. C. *J. fluor. Chem.* **33**, 227 (1986).
4. Banks, R. E. *Fluorocarbons and their Derivatives* 16-17 (Macdonald, London, 1970).
5. Anderson, R. F. & Punderson, J. O. in *Organofluorine Chemicals and their Industrial Applications* (ed. Banks, R. E.) 235-247 (Ellis Horwood, Chichester, 1979).
6. Slinn, D. S. L. & Green, S. W. in *Preparation, Properties, and Industrial Applications of Organofluorine Compounds* (ed. Banks, R. E.) 45-82 (Ellis Horwood, Chichester, 1982).
7. Gething, B., Patrick, C. R., Stacey, M. & Tatlow, J. C. *Nature* **183**, 588-589 (1959).
8. Florin, R. E. in *Fluoropolymers* (ed. Wall, L. A.) 317-380 (Wiley, New York, 1972).
9. *Inclusion Compounds Vols 1-3* (eds Atwood, J. L., Davies, J. E. D. & MacNicol, D. D.) (Academic, London, 1984).
10. Barbour, R. H., Freer, A. A. & MacNicol, D. D. *J. chem. Soc. chem. Commun.* 362-363 (1983).
11. MacNicol, D. D., Mallinson, P. R. & Robertson, C. D. *J. chem. Soc. chem. Commun.* 1649-1651 (1985).
12. Smart, B. E. in *Supplement D: The Chemistry of Halides, Pseudo-Halides and Azides* (eds Patai, S. & Rappaport, Z.) 629-630 (Wiley, Chichester, 1983).
13. Wakselman, C. & Kaziz, C. *J. fluor. Chem.* **33**, 347-358 (1986).
14. Liebman, J. F. *J. fluor. Chem.* **3**, 27-33 (1974).
15. Miller, J. F., Hunt, H. & McBee, E. T. *Analyt. Chem.* **19**, 148-149 (1947).
16. Smart, B. E., Middleton, W. J. & Farnham, W. B. *J. Am. chem. Soc.* **108**, 4905-4907 (1986).
17. Wakselman, C. & Tordeux, M. *J. org. Chem.* **50**, 4047-4051 (1985).
18. Liebman, J. F. *J. fluor. Chem.* **25**, 481-486 (1984).
19. Hardy, A. D. U., MacNicol, D. D. & Wilson, D. R. *J. chem. Soc. Perkin Trans. II* 1011-1019 (1979).
20. Oliver, J. A., Stephens, R., Tatlow, J. C. & Taylor, J. R. *J. fluor. Chem.* **7**, 555-568 (1976).
21. Maruta, M. & Ishikawa, N. *J. fluor. Chem.* **13**, 111-122 (1979).
22. Le Blanc, M. & Reiss, J. G. in *Preparation, Properties and Industrial Applications of Organofluorine Compounds* (ed. Banks, R. E.) 83-138 (Ellis Horwood, Chichester, 1982).
23. Adcock, J. L., Inoue, S. & Lagow, R. J. *J. Am. chem. Soc.* **100**, 1948-1950 (1978).

Anthropogenic chlorofluoromethanes in newly formed Labrador Sea Water

D. W. R. Wallace* & J. R. N. Lazier

Department of Fisheries and Oceans, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, B2Y 4A2, Canada

Convection to 2,000 m in the Labrador Sea during severe winters partially renews Labrador Sea Water (LSW)¹, a water mass traceable at intermediate depth throughout the north Atlantic Ocean (north of 40° N)² and along the western boundary to the Equator³. Anthropogenic chlorofluoromethanes (CFMs) have been used³ to estimate the transit time of LSW from its source region to the Equator, together with its degree of dilution by 'older' waters. CFMs may also be used to constrain 'diagnostic' ocean circulation models⁴. Such applications depend critically on the concentration and the rate of change of CFMs in newly formed LSW. Here we report the first Labrador Sea CFM data, including measurements from a >1,500 m deep column of recently renewed water, that suggest that LSW originated at only 60% saturation with respect to contemporary atmospheric concentrations, in contrast to the 100% usually assumed in interpretations of tracer distributions away from source regions. The newly formed water mass was also colder and less saline than any previously measured LSW. Monitoring of CFMs and other tracers in this region will be essential for establishing boundary conditions for models of the North Atlantic tracer distributions.

The data are from a hydrographic section (Fig. 1) running north-east from the continental shelf and Labrador current to the central Labrador Sea. Potential temperature, θ , against salinity, S , curves (Fig. 2) and vertical sections of potential density anomaly referenced to 1,500 dbar, $\sigma_{1.5}$ (Fig. 3a) and CCl_3F

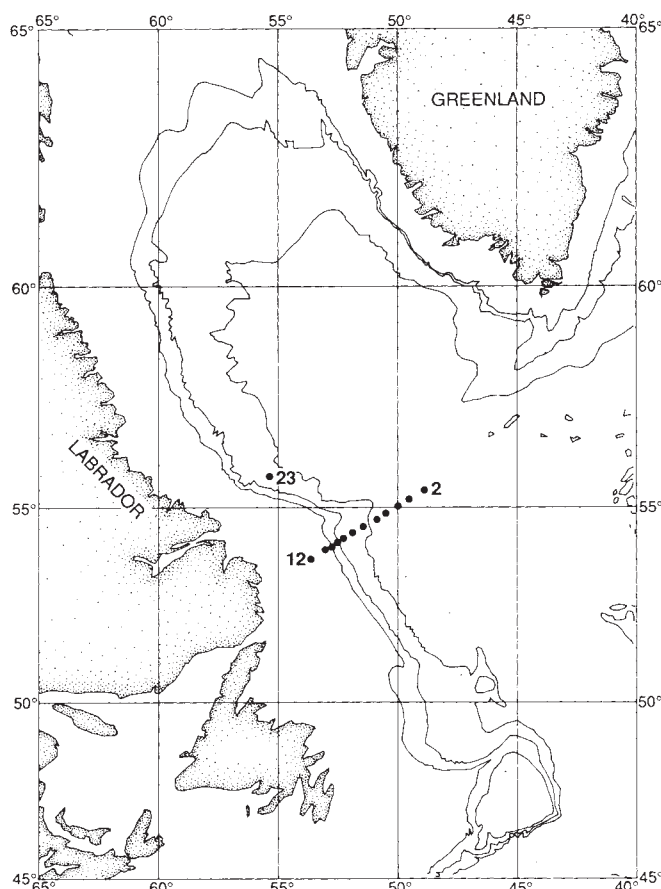


Fig. 1 The Labrador Sea with positions of oceanographic stations occupied during Hudson cruise 86-021, July and August 1986.

(F-11) (Fig. 3b) define and map the three main water masses of the area.

The LSW lies in the approximate interval $34.62 < \sigma_{1.5} < 34.66 \text{ kg m}^{-3}$ and the convective formation process results in characteristic minima in vertical density and CFM gradients (Fig. 4). The LSW is also characterized by minima in salinity and temperature. In 1986, these parameters were 0.07 parts per thousand (p.p.t.) and 0.4°C lower at $\sigma_{1.5} \approx 34.64$ (Fig. 2) than during the years 1969-71. This large shift in properties over 16 years is due partly to a shut-off of deep convection during the late 1960s caused by mild winters and an influx of stabilizing low-salinity water from the Arctic⁵. Limiting convection allowed vertical and horizontal eddy diffusion to increase the temperature and salinity of LSW. In subsequent years, with revived convection to ~2,000 m, the upper water of low salinity was mixed downwards making LSW fresher. The water was also cooled, maintaining the density which tends to remain constant with time. These changes are very large, requiring the addition of 4 m³ of fresh water and the loss of 3.4×10^9 joules of heat from the upper 2,000 m of each square metre of the Labrador Sea showing that this major North Atlantic water mass is strongly affected by decadal-timescale climate variations.

The other two chief water masses in the region lie beneath the LSW. In the interval $34.68 < \sigma_{1.5} < 34.82 \text{ kg m}^{-3}$ is North-East Atlantic Deep Water (NEADW) formed in the eastern basin of the North Atlantic, partly from flow over the Faeroes-Iceland sill. The NEADW is characterized by a salinity maximum and a CFM minimum. The bottom water ($\sigma_{1.5} > 34.82 \text{ kg m}^{-3}$) is North-West Atlantic Bottom Water (NWABW) originating in the Norwegian Sea. This flows as a cyclonic boundary current around the Labrador Sea before entering the western basin of the North Atlantic. This current, the Deep

* Permanent address: Oceanographic Sciences Division, Department of Applied Science, Brookhaven National Laboratory, Upton, New York 11973, USA.

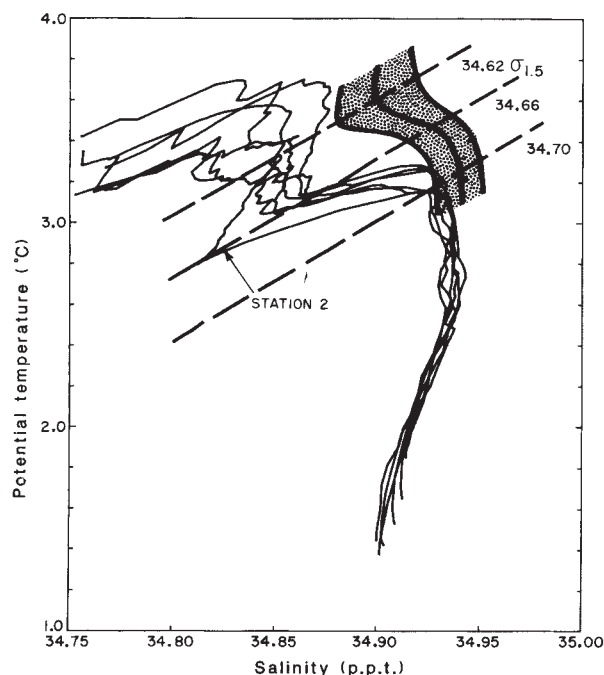


Fig. 2 Potential temperature against salinity curves for deep Labrador Sea stations occupied during 1986. Water of exceptionally low temperature and salinity near $\sigma_{1.5} = 34.66$, Station 2, is indicative of renewal by deep convection, probably during the previous winter. The shaded area outlines the means and standard deviations of observations obtained between 1969 and 1971 at Ocean Weather Station Bravo in the central Labrador Sea¹², when deep convection did not occur.

Western Boundary Undercurrent, has very high CFM content ($\sim 2.5 \text{ pmol kg}^{-1}$ (F-11) and $\sim 1.1 \text{ pmol kg}^{-1}$ (F-12)).

Station 2 at the end of the section exhibits distinctive features which suggest that LSW has been recently renewed by deep convection. The vertical separation between the $\sigma_{1.5}$ surfaces 34.62 and 34.66 is 1,500 m, indicative of a very small vertical density gradient. Historical data (121 stations, 1965–86) from the central Labrador Sea, for comparison, give an average separation of 705 m (± 250 m). The largest vertical separation measured, of 2,000 m, was observed during severe winter conditions in a region of active convection¹. Another indication of recent renewal is the anomalous temperature and salinity of this layer ($\theta = 2.80^\circ\text{C}$, $S = 34.10$ p.p.t., see Fig. 2): up to 0.3°C colder and 0.04 p.p.t. less saline than water of the same density at the other stations. These layer characteristics represent the most extreme variant of LSW observed to date, being 0.09°C colder and 0.033 p.p.t. less saline than observed during a convection event in 1976^{1,6}. Finally, the concentrations of two CFMs, CCl_3F (F-11) and CCl_2F_2 (F-12) are nearly constant over a depth interval of at least 1,600 m (Fig. 4) and have higher concentrations in water with $\sigma_{1.5} \approx 34.66$ than are found elsewhere in the section (F-11 is $0.5\text{--}1 \text{ pmol kg}^{-1}$ higher). Because atmospheric CFM concentrations have been increasing with time, water that has recently been in contact with the atmosphere should have the highest CFM levels.

The convectively formed feature at Station 2 was not present at Station 3, 45 km distant. The anomalous hydrographic structure probably had a horizontal length scale comparable to the deformation radius (~ 8 km for a 1,500 m deep water column), similar to eddy-scale features previously observed in the Labrador^{1,7} and Weddell⁸ seas. Our observations suggest that such features have long lifetimes (at least 3–4 months).

The CFM concentrations in the deep mixed layer at Station 2 should correspond closely to those achieved immediately after convection, because horizontal mixing with the surroundings is

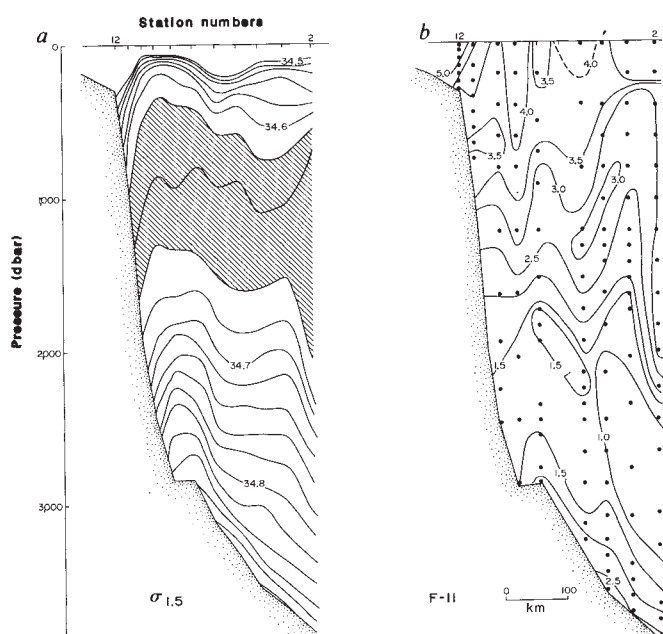


Fig. 3. *a*, Vertical section of $\sigma_{1.5}$ from stations 2 to 12. The shaded area delimits Labrador Sea Water. *b*, Vertical section of the anthropogenic chlorofluoromethane CCl_3F (F-11) from stations 2 to 12.

evidently restricted (Fig. 2). It is important to note that CFM levels lie at 60% saturation with respect to contemporary atmospheric concentrations; far below the 100% saturation previously assumed in models of the circulation of this water mass³. This implies that deep convection took place too rapidly for air-sea gas exchange to bring CFM levels to near equilibrium.

This has recently been shown to be the case for dissolved oxygen during winter convection in the Labrador Sea⁶, and we have modified a convection model^{1,6} to examine the degree of equilibration of CFMs that a mixed layer can attain during deep convection. The model removes heat and water vapour ($0.003 \text{ m per day}^{1,6}$) from a mixed layer, and allows it to deepen to maintain static stability in relation to the underlying water column. Exchange of gases with the atmosphere is parameterized using a piston velocity of 15.6 m per day , a reasonable estimate⁹ for winter conditions in this region, and published CFM solubility data¹⁰. The model was run separately using vertical profiles from each of the deep stations as the initial condition, and atmospheric CFM mixing ratios measured in the region during July–August 1986: 422×10^{-12} (F-12) and 233×10^{-12} (F-11). The model was run for 240 days to simulate the period August to March.

Model results suggested that: (1) most of the vertical profiles studied were not prone to convection below 1,000 m under 'normal' conditions (mean heat flux¹¹ $\sim 150 \text{ W m}^{-2}$). (2) Deep convection (to 1,400 m) could be driven by such a heat flux only at Station 7. The resulting final CFM gas saturations were $\sim 88\%$. Even with no gas exchange (but accounting for effects of temperature and salinity changes on gas solubility), the final saturations were $\sim 65\%$, which remains higher than observed at Station 2. (3) Forcing the other water columns with extremely high and probably unrealistic sustained heat fluxes (250 W m^{-2}) did result in deep mixing, but final saturations, with gas exchange, were $\sim 80\%$, compared with $55\text{--}65\%$ without gas exchange. These results suggest that some form of preconditioning of a water column, for example a cyclonic circulation⁷, might be required for deep convection to occur. The apparent overestimation of post-convective CFM saturations probably results from a lack of LSW renewal for some years before 1986 which would have caused pre-convection CFM concentrations to have

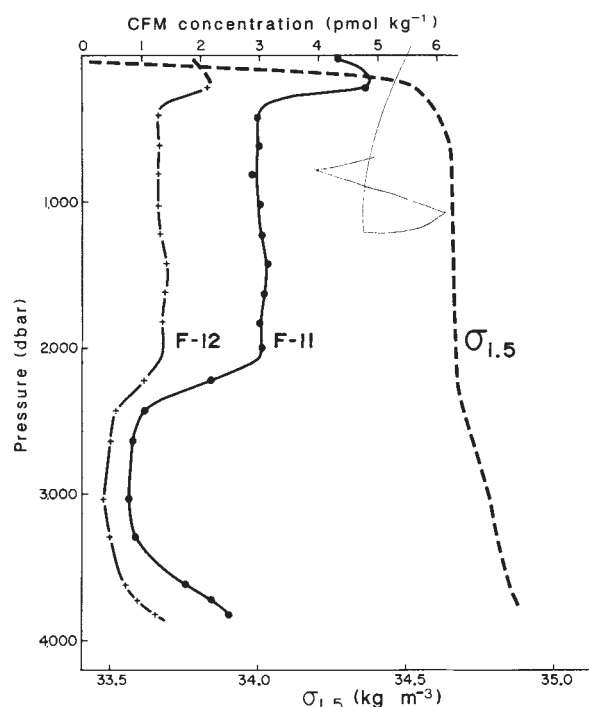


Fig. 4 Vertical profiles of two chlorofluoromethanes CCl_2F_2 (F-12) and CCl_3F (F-11) and $\sigma_{1.5}$ at station 2. A convectively formed deep mixed layer has been covered by an $\sim 500\text{-m}$ -deep layer of water, more typical of the near-surface, southern Labrador Sea. CFMs were measured using purge-and-trap gas chromatography with electron-capture detection¹³. Analytical precision for these data was the greater of: 1% or $0.016\text{ pmol kg}^{-1}$ for F-12, and 1.5% or $0.042\text{ pmol kg}^{-1}$ for F-11. CFM concentrations are reported relative to the SIO (Scripps Institution of Oceanography) 1986 calibration scale (R. F. Weiss, personal communication).

been lower than the initial conditions used with the models. An air-sea CFM flux comparable with the modelled flux could therefore have occurred during the 1985–86 winter, but sufficient to result in only 60% saturation because of the lower initial concentrations. Simulations with the model suggest that the pre-convection inventory of CFM in the upper 2,000 m of the water column at Station 2 might have been only 60–70% of the amount observed there in July–August 1986. In any case, the model results confirm that LSW in its source region does not necessarily attain 100% saturation with respect to CFMs during one year or even two successive years of deep convective renewal.

The concentrations of CFMs in LSW varies both between and within stations. Defining LSW as before, F-11 concentrations ranged from 2.1 to 3.7 pmol kg^{-1} . The mean depth-averaged concentrations in LSW were: 3.05 pmol kg^{-1} (F-11) and 1.32 pmol kg^{-1} (F-12). These values are almost identical to those observed in the deep mixed layer at Station 2 (Fig. 4), and should therefore provide an accurate representation of the initial conditions in LSW during 1986. Use of CFM data as a constraint for ocean circulation models⁴ however, will require time series of the CFM content of source waters. Because newly formed LSW has been observed with significantly undersaturated CFM levels, and because the formation process is highly variable in time^{1,5} and difficult to model, it seems clear that a CFM source function for LSW cannot be estimated accurately at present. In addition, the historically extreme LSW temperature and salinity characteristics reported here emphasize that even these characteristics are highly variable in time. It is therefore essential that long-term monitoring of CFMs and other tracers be initiated in this region in order that the information contained in their distributions can be exploited.

This research was funded by the Government of Canada, Department of Fisheries and Oceans.

Received 24 August 1987; accepted 4 January 1988.

1. Clarke, R. A. & Gascard, J.-C. *J. phys. Oceanogr.* **13**, 1764–1778.
2. Talley, L. D. & McCartney, M. S. *J. phys. Oceanogr.* **12**, 1189–1205 (1982).
3. Weiss, R. F., Bullister, J. L., Gammon, R. H. & Warner, M. J. *Nature* **314**, 608–610 (1985).
4. Woods, J. D. *Nature* **314**, 501–511 (1985).
5. Lazier, J. R. N. *Atmos.-Ocean* **18**, 227–238 (1981).
6. Clarke, R. A. & Coote, A. R. *J. phys. Oceanogr.* (in the press).
7. Gascard, J.-C. & Clarke, R. A. *J. phys. Oceanogr.* **13**, 1779–1797 (1983).
8. Gordon, A. L. *J. phys. Oceanogr.* **8**, 600–612 (1978).
9. Roether, W. in *Dynamic Processes in the Chemistry of the Upper Ocean* (eds Burton, J. D., Brewer, P. G. & Chesselet, R.) 117–128 (Plenum, New York, 1986).
10. Warner, M. J. & Weiss, R. F. *Deep Sea Res.* **32**, 1485–1497 (1985).
11. Smith, S. D. & Dobson, F. W. *Atmos.-Ocean* **22**, 1–22 (1984).
12. Shuh, J. L. *Oceanographic Report No. CG373-78*, Oceanographic Observations North Atlantic Ocean Station Bravo, Terminal Report 1964–1974 (US Coast Guard Oceanography Unit, Washington DC, 1978).
13. Wallace, D. W. R. & Moore, R. M. *J. geophys. Res.* **90**, 1155–1166 (1985).

Mass extinctions, atmospheric sulphur and climatic warming at the K/T boundary

Michael R. Rampino*[†] & Tyler Volk*

* Department of Applied Science, New York University, New York, New York 10003, USA

[†] NASA, Goddard Space Flight Center, Institute for Space Studies, New York, New York 10025, USA

A connection has recently been proposed between cloud albedo over the oceans and the release of dimethyl sulphide (DMS) by marine algae. DMS acts as a precursor for most of the cloud condensation nuclei (CCN) in the marine atmosphere¹. The mass extinctions at the Cretaceous/Tertiary (K/T) boundary include about 90% of marine calcareous nannoplankton^{2,3}, and carbon isotope data show that marine primary productivity as a whole was drastically reduced for at least several tens of thousands of years, and perhaps up to a million years after the extinction event^{4–6}. The elimination of most marine calcareous phytoplankton would have meant a severe decrease in DMS production, leading to a drastic reduction in CCN and hence marine cloud albedo. Here we examine the possible climatic effects of a drastic decrease in CCN associated with a severe reduction in the global marine phytoplankton abundance. Calculations suggest that a reduction in CCN of more than 80%, and the resulting decrease in marine cloud albedo, could have produced a rapid global warming of 6°C or more. Oxygen isotope analyses of marine sediments from many parts of the world have been interpreted as indicating a marked warming coincident with the demise of calcareous nannoplankton at the K/T boundary. Decreased marine cloud albedo, and resulting high sea surface temperatures could have been a factor in the maintenance of low productivity in the 'Strangelove Ocean' period following the K/T extinctions.

Charlson *et al.*¹, and subsequent work⁷, have shown that the precursor for most CCN over the oceans is the DMS released by marine phytoplankton, which oxidizes to form a sulphate aerosol. If the total liquid water content in the clouds is held constant, changes in the number density of CCN would affect the cloud droplet size, which affects the cloud albedo. Their model¹ has quantified the effects of either increasing the number density of cloud droplets, which increases the cloud albedo, or decreasing the number density with its associated decrease in the cloud albedo.

We use the model results from Charlson *et al.* (their Fig. 1), which show the effect upon stratiform top-of-cloud albedo due to variations in the CCN number density. Following their example, these albedo variations have been converted to changes in solar albedo at the top of the atmosphere averaged over the

Earth's surface, and then into equivalent variations in the solar constant. The range spanned by these equivalent solar constant variations ($1.0 \leq S/S_0 \leq 1.06$) corresponds to the range studied by Wetherald and Manabe⁸ with a general circulation model (GCM). We therefore apply their results to convert the equivalent solar constant variations into associated variations in global average surface temperature (ΔT). Figure 1 shows these variations as a function of normalized DMS concentration for $0.1 \leq [\text{DMS}]/[\text{DMS}]_{\text{ref}} \leq 1.0$. (It was not possible to extrapolate the results of Charlson *et al.* to values < 0.1 , but these would give somewhat higher temperatures.) Figure 1 shows the results for reference top-of-cloud albedos of 0.3, 0.5 and 0.7, which span the range of satellite-derived albedos over marine stratus and stratocumulus¹.

The GCM results of Wetherald and Manabe⁸ give a climate sensitivity (the slope of the $\Delta T(S/S_0)$ curve) of 2.2 °C per 1% change in the solar constant when $S = S_0$, which is higher than the value of 1.85 (from an earlier paper by Wetherald and Manabe⁹) applied by Charlson *et al.*, but consistent with the value of 2.0 found by another GCM study of Hansen *et al.*¹⁰. Also, the climate sensitivity of Wetherald and Manabe is not constant, but smoothly decreases to 0.9 when $S = 1.06S_0$. Climate sensitivity was found to behave differently in a GCM study of the Cretaceous¹¹, in which the climate sensitivity to a quadrupling of CO_2 was actually higher by 33% in the Cretaceous than in the present-day model. This increase in sensitivity is probably the result of the greater importance of water vapour climate feedback in the Cretaceous experiment, due to higher reference temperatures and the nonlinear relationship between surface temperature and saturation vapour pressure. Here we use the Wetherald and Manabe⁸ results as a reasonable method for converting solar constant variations to global temperatures, and note that even if the temperatures associated with the smallest values of $[\text{DMS}]/[\text{DMS}]_{\text{ref}}$ in Fig. 1 were decreased by several °C due to possible negative feedbacks^{1,7}, our conclusions would remain essentially the same.

Figure 1 shows that for DMS concentrations $< 20\%$ of the reference value, global temperatures could rise by ≥ 6 °C, and, depending upon top-of-cloud albedo, by nearly 10 °C when DMS is reduced to 10% of the reference value. Even a 50% reduction in DMS would produce a significant global warming of ~ 3 –4 °C. The thermal mass and mixing structure of the ocean would delay the full surface warming by only several thousand years. These results indicate that a drastic reduction in DMS production, if maintained, could lead to a substantially warmer Earth essentially simultaneous with a phytoplankton extinction event.

Two separate pieces of evidence point to a trauma in the global marine biosphere 66 million years ago, at the K/T boundary. First, a number of studies report a negative anomaly in $\delta^{13}\text{C}$ of up to 3‰ in the carbonate fine fraction ($< 63 \mu\text{m}$, largely nannoplankton debris), and in individual planktonic foraminifera, in K/T boundary sediments. This anomaly appears to be a worldwide phenomenon^{12–18}, and suggests a surface ocean of very low primary productivity, the so-called 'Strangelove Ocean'^{16,18}. In the South Atlantic Deep Sea Drilling Project (DSDP) Site 524, the carbon isotope anomaly begins just above the K/T boundary (as marked by an iridium anomaly in a thin boundary clay) and reaches the minimum value $\sim 40,000$ –50,000 years later⁴; $\delta^{13}\text{C}$ values did not return to their pre-boundary values for at least 300,000–400,000 years. The magnitude of the $\delta^{13}\text{C}$ decrease varies from locality to locality, and it has been detected solely in surface-water calcareous organisms. Arthur *et al.*⁶ suggest that the significant reduction in surface-ocean primary productivity lasted more than 1 Myr after the K/T boundary.

Second, the K/T boundary is marked by a drastic reduction in CaCO_3 deposition in the deep sea⁶ reflecting mainly the mass extinction of calcareous nannoplankton and foraminifera. Calcium carbonate deposition was reduced for at least 350,000

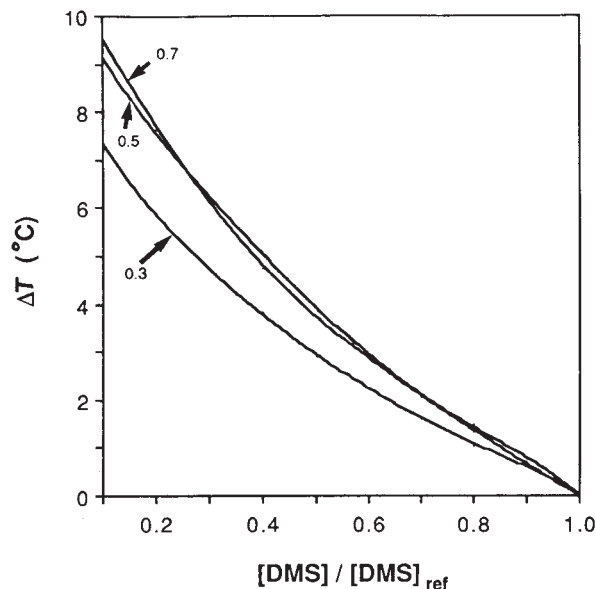


Fig. 1 Change of globally averaged surface temperature as a function of normalized DMS concentration, assumed equivalent to global DMS production rate by phytoplankton. These results were obtained from the model of Charlson *et al.*¹ relating the visible albedo at the top of the reference cloud to the number density of cloud droplets (assumed equivalent to DMS concentration) (see Fig. 1 in ref. 1). Their model used a reference cloud droplet radius of 8 μm , but they stated that the results are still valid within 20% for droplet radii within the range found in real clouds (4–500 μm). Then, following the technique shown in their Table 1, the changes in top-of-cloud albedo for oceanic stratiform water clouds were converted to equivalent changes in the solar constant by multiplying sequentially by factors of 0.81 (to convert top-of-cloud albedo to solar albedo at the top of the atmosphere), 0.31 (to convert from the area covered by oceanic stratiform water clouds to an average over the Earth's surface area) and $1/(1 - \text{present planetary albedo}) = 1.43$ (to convert from the planetary albedo change to equivalent change in the solar constant, where the present planetary albedo is taken as 0.3). Finally, we used the GCM model results of Wetherald and Manabe⁸ (see their Table 8 and Fig. 20 for the model with variable clouds) to calculate the global average surface temperatures for the various values of the equivalent solar constant. Curves are shown for three different values (0.3, 0.5 and 0.7) of the visible albedo at the top of the reference cloud.

years⁴, and in some cases for as long as 1 Myr⁶. According to Zachos and Arthur¹⁸, this decrease in carbonate was the result of a severe decrease in surface primary production, not a long-term dissolution event.

The DMS production per unit productivity seems to vary significantly among species¹⁹ (M. O. Andreae, personal communication). Thus, the general correspondence between DMS and primary productivity (for example, relatively high DMS in upwelling regions¹⁹) might be due to a roughly constant ecological mix of high- and low-DMS producers. DMS flux to the atmosphere has been shown to be correlated seasonally and regionally with solar flux⁷, but solar flux also drives productivity, especially seasonally in high latitudes. The inference of a decrease in DMS associated with a global collapse of primary productivity as shown by the $\delta^{13}\text{C}$ data across the K/T boundary requires only the reasonable assumption of some regularity in the mix of high- and low-DMS producers on both sides of the boundary. However, because Andreae (personal communication) has proposed that the calcareous nannoplankton are one of the main or perhaps even the dominant producers of DMS, the mass extinction of this group at the boundary is strong evidence for a significant change in DMS flux.

Is the predicted surface warming observed? Oxygen isotope

analyses across the K/T boundary give differing results. A number of analyses suggest a possible negative excursion of $\leq 2.5\%$ in $\delta^{18}\text{O}$ at or just above the K/T boundary, which in some cases is coincident with the most negative $\delta^{13}\text{C}$ anomaly in the boundary sections^{4,15,17,20-22}. Assuming that a 0.2% decrease in $\delta^{18}\text{O}$ represents a 1°C temperature increase, this negative $\delta^{18}\text{O}$ excursion has been interpreted as possibly indicating a dramatic warming of ocean surface waters by as much as 10–12°C (refs 4, 5, 15, 20, 21).

For example, boundary sections in Spain at Caravaca, Sopelana and Zumaya show maximum negative $\delta^{18}\text{O}$ excursions in bulk carbonate samples, at or just above the K/T boundary, ranging from -1.4 to -2.4% , corresponding to warmings of ~ 7 – 12°C , closely correlated with the maximum negative $\delta^{13}\text{C}$ excursion^{15,21,23}. Similarly, bulk carbonate in the section at Biarritz, France shows a maximum negative $\delta^{18}\text{O}$ excursion of $\sim 1.5\%$ across the boundary, whereas the boundary sequence at Lattegebirge in southwestern Germany, although possibly affected by diagenesis, has a possible maximum negative $\delta^{18}\text{O}$ departure of -2.5% , in both cases corresponding to the maximum negative excursion in $\delta^{13}\text{C}$ ¹⁵. In the South Atlantic, at DSDP Site 524, Hsü *et al.*⁴ report an average $\delta^{18}\text{O}$ excursion in the carbonate fine fraction of -0.4% between pre- and post-boundary samples, with a difference of -2% between extremes in the standard deviation, suggesting warmings of at least 2°C and possibly up to 10°C . The maximum negative excursion in $\delta^{18}\text{O}$ occurs at about the same level as the most negative $\delta^{13}\text{C}$ anomaly, in the Lower Palaeocene⁴. A negative departure of $\sim 1\%$ in $\delta^{18}\text{O}$ in benthic foraminifera, in some sections, at about the same time indicates possible warming of ocean bottom waters of $\leq 5^\circ\text{C}$ over late Cretaceous values^{5,20}. In general, the inferred sea surface temperature increases seem to be greater in European and Atlantic sites than in the Pacific^{15,18}.

In contrast, other studies of K/T boundary sections have not identified a marked negative $\delta^{18}\text{O}$ departure^{18,23,24}, and the $\delta^{18}\text{O}$ anomaly in some sections has been interpreted as a selective diagenetic overprint^{18,22}. One would expect, however, that the magnitude of the temperature increase, and hence the oxygen isotope anomaly, might vary considerably from site to site, as seen in the record. The increase in surface water temperatures might be suppressed in the tropics, if most of the increased solar energy received there goes into evaporation, and not into sensible heating²⁵. It may be significant, therefore, that the evidence for little or no increase in sea surface temperatures comes from the sections at the lowest palaeolatitudes, DSDP sites 47.2 and 577 at $\sim 20^\circ\text{N}$ palaeolatitude, and DSDP site 465 at $\sim 15^\circ\text{N}$ palaeolatitude, in the Pacific. It should also be recognized that the early Danian seems to have been a time of unstable and fluctuating oceanographic and climatic conditions^{4,15,18}, and some sections perhaps contain hiatuses. For example, several of the DSDP sites studied by Zachos and Arthur¹⁸ have no measurable iridium anomaly, suggesting the possible incompleteness of these boundary sections.

The occurrence of a significant negative $\delta^{18}\text{O}$ anomaly in boundary sections from various parts of the world, and the coincidence of the oxygen isotope anomaly with the marked negative $\delta^{13}\text{C}$ anomaly in a number of boundary sections^{4,15,21,22} suggest that the oxygen isotope anomaly is not a local diagenetic feature. Diagenesis cannot be ruled out as a factor in the apparent $\delta^{18}\text{O}$ fluctuations in some sections, and further isotope analyses on sequences with high sedimentation rates, shallow burial and no apparent hiatuses at the K/T boundary are needed to resolve the problem¹⁸. Deep Sea Drilling sites in relatively high palaeolatitudes might be the best sections for oxygen isotope analyses designed to test further for a significant warming in sea surface temperatures at the K/T boundary (J. Zachos, personal communication).

Several authors have suggested that a greenhouse warming from increased atmospheric CO_2 and/or H_2O could have taken place at the K/T boundary²⁶⁻²⁸, and Kasting *et al.*²⁹ have model-

led the possible global temperature increase from elevated atmospheric CO_2 following an impact event. We believe that the considerable decrease in marine cloud albedo predicted here could have provided an independent mechanism for climate warming, which could significantly enhance any possible warming resulting from increases in atmospheric greenhouse gases. The model proposed here could be further tested with global climate model sensitivity studies of the effects of changes in marine cloud albedo on sea temperatures using reconstructed late Cretaceous boundary conditions^{11,30}. The available evidence suggests that the extinction of virtually all late Cretaceous calcareous nannoplankton and, as inferred from the $\delta^{13}\text{C}$ record, the drop in global productivity by phytoplankton, in general, could have precipitated a severe warming. This heating, and its destabilizing effects on global climate, combined with the proposed impact-induced³¹, and possibly related volcanic^{32,33} trauma at the boundary could have inhibited the full recovery of the marine biosphere, and may have contributed to the initiation and maintenance of Strangelove Ocean conditions for the observed period of $\geq 10^5$ years.

Similar carbon isotope anomalies have been reported at other major geological boundaries; the Precambrian/Cambrian³⁴⁻³⁷, the Devonian/Carboniferous³⁸ and the Permian/Triassic³⁹ (although their significance has been questioned^{40,41}). It may be, therefore, that episodes of reduced marine cloud albedo and related global warmings accompanied other mass extinctions in the geological record, and these may be detectable through careful palaeotemperature analyses at these boundaries.

We thank K. Caldeira, M. I. Hoffert, R. B. Stothers, S. Warren and J. C. Zachos for helpful discussions, and M. O. Andreae for a review. Research support for M.R.R. was provided by the Center for Global Habitability at Columbia University. T.V. was supported by NASA.

Received 21 September 1987; accepted 7 January 1988.

1. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).
2. Thierstein, H. R. *Spec. Publ. Soc. Econ. Miner. Petrol.* **32**, 355–394 (1981).
3. Thierstein, H. R. *Spec. Pap. geol. Soc. Am.* **190**, 385–399 (1982).
4. Hsü, K. J. *et al. Science* **216**, 249–256 (1982).
5. Hsü, K. J., McKenzie, J. A. & He, Q. X. *Spec. Pap. geol. Soc. Am.* **190**, 317–328 (1982).
6. Arthur, M. A., Zachos, J. C. & Jones, D. S. *Cret. Res.* **8**, 43–54 (1987).
7. Bates, T. S., Charlson, R. J. & Gammon, R. H. *Nature* **329**, 319–321 (1987).
8. Wetherald, R. T. & Manabe, S. *J. Atmos. Sci.* **37**, 1485–1510 (1980).
9. Wetherald, R. T. & Manabe, S. *J. Atmos. Sci.* **32**, 2044–2059 (1975).
10. Hansen, J. E. *et al. in Climate Processes and Climate Sensitivity* (eds Hansen, J. E. & Takahashi, T.) 130–163 (Am. Geophys. Union, Washington, DC, 1984).
11. Barron, E. J. & Washington, W. M. in *The Carbon Cycle and Atmospheric CO_2 : Natural Variations Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 546–553 (Am. Geophys. Union, Washington, DC, 1985).
12. Brenneke, J. L. & Anderson, T. F. *Am. geophys. Union Trans.* **58**, 415 (1977).
13. Thierstein, H. R. & Berger, W. H. *Nature* **276**, 461–466 (1978).
14. Scholle, P. A. & Arthur, M. A. *Am. Ass. Petrol. Geol. Bull.* **64**, 67–87 (1980).
15. Perch-Nielsen, K., McKenzie, J. A. & He, Q. X. *Spec. Pap. geol. Soc. Am.* **190**, 353–371 (1982).
16. Hsü, K. J. & McKenzie, J. A. in *The Carbon Cycle and Atmospheric CO_2 : Natural Variations Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 487–492 (Am. Geophys. Union, Washington, DC, 1985).
17. Hsü, K. J. in *Mesozoic and Cenozoic Oceans* (ed. Hsu, K. J.) 75–84 (Am. Geophys. Union, Washington, DC, 1986).
18. Zachos, J. C. & Arthur, M. A. *Paleoceanogr.* **1**, 5–26 (1986).
19. Andreae, M. O. in *The Role of Air–Sea Exchange in Geochemical Cycling* (ed. Buat-Ménard, P.) 331–362 (Reidel, Dordrecht, 1986).
20. Boersma, A. *et al. Init. Rep. DSDP* **43**, 695–718 (1979).
21. Smit, J. *Spec. Pap. geol. Soc. Am.* **190**, 329–352 (1982).
22. Margolis, S. V. *et al. Paleoceanogr.* **2**, 361–377 (1987).
23. Boersma, A. & Shackleton, N. J. *Init. Rep. DSDP* **62**, 513–526 (1981).
24. Zachos, J. C. *et al. Init. Rep. DSDP* **86**, 513–532 (1985).
25. Hoffert, M. I. *et al. J. Atmos. Sci.* **40**, 1659–1668 (1983).
26. McLean, D. M. *Science* **201**, 401–406 (1978).
27. McLean, D. M. *Cret. Res.* **6**, 235–259 (1985).
28. Emiliani, C., Kraus, E. B. & Shoemaker, E. M. *Earth planet. Sci. Lett.* **55**, 317–334 (1981).
29. Kasting, J. F., Richardson, S. M., Pollack, J. B. & Toon, O. B. *Am. J. Sci.* **286**, 361–389 (1986).
30. Glancy, T. J. Jr, Barron, E. J. & Arthur, M. A. *Paleoceanogr.* **1**, 523–537 (1986).
31. Alvarez, W. *Eos* **67**, 649–658 (1986).
32. Courtillot, V. *et al. Earth planet. Sci. Lett.* **80**, 361–374 (1986).
33. Rampino, M. R. *Nature* **327**, 468 (1987).
34. Hsü, K. J. *et al. Nature* **316**, 809–811 (1985).
35. Margaritz, M. *et al. Nature* **320**, 258–259 (1986).
36. Tucker, M. E. *Nature* **319**, 48–50 (1986).
37. Aharon, P., Schidlowski, M. & Singh, I. B. *Nature* **327**, 699–702 (1987).
38. Sun Yijun *et al. in Contr. 27th Int. Geol. Congr.* 225–234 (Science Press, Beijing, 1984).
39. Dao-Yi, X. *et al. Nature* **321**, 854–855 (1986).
40. Awramik, S. M. *Nature* **319**, 696 (1986).
41. Morris, S. C. & Bengtson, S. *Nature* **319**, 696–697 (1986).

Constraints on granulite genesis from carbon isotope compositions of cordierite and graphite

J. Vry, P. E. Brown, J. W. Valley & J. Morrison

Department of Geology and Geophysics, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

Carbon dioxide is a common constituent in the channels of high-grade metamorphic cordierites¹⁻⁶. Here we show that analysis of the stable isotopes of carbon in cordierite (crd) or graphite (gr) can provide a means to discriminate between two controversial models of granulite formation: (1) large-scale CO₂ infiltration⁷, or (2) magmatic processes⁸⁻¹¹, which might result in fluid-absent metamorphism. The fractionation of carbon isotopes between cordierites and graphites in high-grade rocks approximates the predicted equilibrium value. This suggests that cordierites can preserve peak metamorphic carbon isotope and channel gas compositions. Isotopically light carbon values in many granulite-facies cordierites ($\delta^{13}\text{C}_{\text{crd}} < -15$) and graphites severely restrict theories of massive CO₂ infiltration ($\delta^{13}\text{C}_{\text{fluid}} \approx -7$). CO₂/rock ratios from a number of granulite terranes are estimated to be ≤ 0.01 by weight (~ 0.016 on a molar oxygen basis), unless the rocks contain substantial additional carbon. These data emphasize the importance of local fluid buffering in the granulite facies, consistent with dehydration by partial melts.

Cordierite is of interest as a potential monitor for important volatile species such as H₂O and CO₂ in metamorphic rocks. Literature analyses of granulite-facies cordierites^{4,6} commonly show relatively high contents of channel CO₂ (up to 2.2 wt %)⁴. Such high CO₂ contents have been interpreted^{4,7} on the basis of fluid-saturated experiments¹² to reflect high molar ratios and partial pressures of CO₂ in granulite-facies fluids. Uncertainties have remained, however, as to whether measured cordierite channel gas compositions in fact represent equilibration with a peak metamorphic fluid phase, or instead, the effects of leakage and reequilibration. Such questions pertain directly to the current controversy as to the relative importance of magmatic processes⁸⁻¹¹ and fluid infiltration⁷ in stabilization of the deep continental crust, and can be examined using stable isotope measurements¹³.

We have analysed the carbon isotope ratio of cordierite channel CO₂ in eighteen samples from nine different metamorphic terranes, and of coexisting, coarsely crystalline graphite wherever possible (seven samples, graphite flakes up to 2 mm in size; see Table 1). Cordierite channel CO₂ was extracted for carbon isotopic analysis by heating pure, gemmy cordierite (200–775 mg) with excess CuO at 1,050 °C in a vacuum extraction line. Hand picking required 5–30 h per sample; great care was exercised to avoid graphite, sulphides or any regions of possible alteration. Carbon isotope ratios were determined with a Finnigan MAT 251 mass spectrometer, and are expressed in conventional per mil notation ($\delta^{13}\text{C}$) relative to the PDB standard.

The question of whether cordierite channel gas compositions reflect equilibration with a peak metamorphic fluid can be examined in samples that contain coexisting cordierite and graphite. The isotopically distinctive graphites from the Pikwitonei granulite domain (PGD) in Manitoba, Canada, are interpreted to represent original Archaean organic carbon values¹⁴, and thus provide a useful marker. In the PGD, four cordierites have $\delta^{13}\text{C} = -31.8$ to -34.9 , coexisting graphites have $\delta^{13}\text{C} = -36.1$ to -41.8 , and the values for the fractionation between cordierite and graphite, $\Delta^{13}\text{C}_{\text{crd-gr}}$, are 4.3, 6.6, 7.0 and 7.7¹⁵. Similar values of Δ (3.3, 4.1 and 3.1) were obtained for samples from three other amphibolite- or granulite-facies terranes (Table 1). Because CO₂ molecules in cordierite occupy

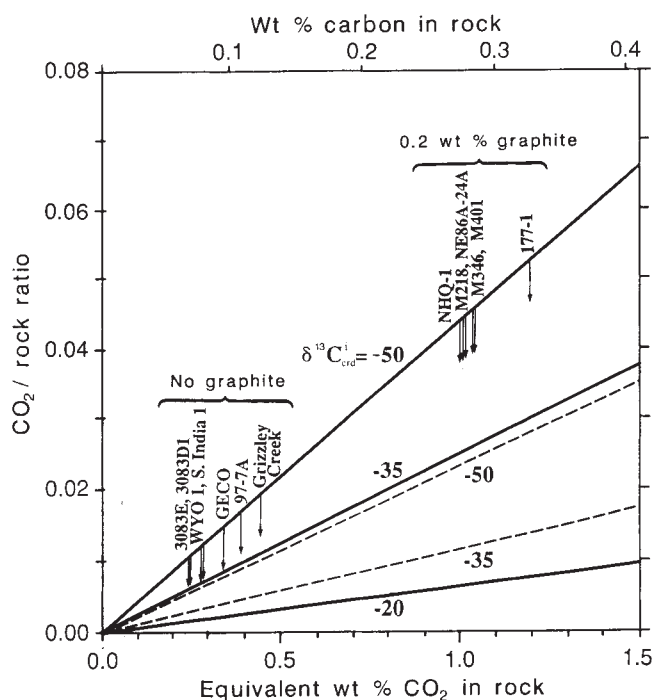


Fig. 1 Maximum values of CO₂/rock ratio (by weight) that are consistent with the preservation of premetamorphic low $\delta^{13}\text{C}$ values in cordierite-bearing rocks, from equation (1). Values of initial $\delta^{13}\text{C}_{\text{crd}}$ are -50 , -35 and -20 , with solid isopleths for final $\delta^{13}\text{C}_{\text{crd}} = -15$ and dashed isopleths for final $\delta^{13}\text{C}_{\text{crd}} = -20$. Initial $\delta^{13}\text{C}_{\text{fluid}}$ is -7 , consistent with a mantle origin; higher values from marble result in lower CO₂/rock ratios. This figure may also be used for graphite samples by applying $\Delta_{\text{gr-crd}} = -5.6$ to -8.2 at 600–800 °C. For the ten granulite-facies samples in Table 1 with measured $\delta^{13}\text{C}_{\text{crd}} \leq -15$, only very small amounts of CO₂ exchange can have occurred (CO₂/rock < 0.06). Larger amounts of mantle- or marble-derived CO₂ would have raised $\delta^{13}\text{C}_{\text{crd}}$ to above -15 . Thus, low values of $\delta^{13}\text{C}_{\text{crd}}$ or $\delta^{13}\text{C}_{\text{gr}}$ rule out infiltration by large amounts of mantle CO₂, consistent with partial melting as the cause of granulite-facies dehydration.

sites in channels, rather than strongly bonded sites in the mineral framework, the fractionation between channel CO₂ and fluid CO₂ is probably small. The ^{13}C fractionation in cordierite in equilibrium with graphite should therefore approximate $\Delta^{13}\text{C}_{\text{CO}_2\text{-gr}}$, which at granulite-facies temperatures is 5.6–8.2^{16,17}. The measured values of $\Delta^{13}\text{C}_{\text{crd-gr}}$ from metamorphic pairs approximate the predicted fractionation and show no reversals of Δ . The apparent preservation of peak metamorphic channel gas compositions is further supported by the unusually low $\delta^{13}\text{C}$ of PGD samples, as no other source of such low values is known.

Peak metamorphic CO₂ contents are frequently preserved in cordierite, even though the loosely bonded channel gas molecules might be expected to undergo post-metamorphic exchange. This conclusion is supported by experimental dehydration studies¹⁸, which suggest that relatively large contents of water (0.5–0.6 mol per formula unit) may persist metastably in cordierite over geological periods. It is possible that bonded alkali cations block the channels, even at the low alkali content of many granulite cordierites (≤ 0.03 atoms per formula unit; J.K.V., unpublished data), or that the large CO₂ molecules themselves block the channel.

If cordierites reflect equilibrium with metamorphic fluids, then any low $\delta^{13}\text{C}_{\text{crd}}$ proves that large amounts of deep-seated CO₂ ($\delta^{13}\text{C}_{\text{fluid}} \approx -7$) have not infiltrated the rock. Most determinations of 'mantle' carbon $\delta^{13}\text{C}$ are between -5 and -10 , although the mantle may not be completely homogeneous in isotopic composition ($\delta^{13}\text{C}$ values for some diamonds range from 0 to -33 , but average approximately -6 (ref. 19)). The isotopic composition of CO₂ in many magmas is near -5 (ref. 19).

Table 1 Carbon isotope compositions of metamorphic cordierites and graphites

Sample	$\delta^{13}\text{C}_{\text{crd}}$ (‰)	$\delta^{13}\text{C}_{\text{gr}}$ (‰)	$\Delta_{\text{crd-gr}}$ (‰)	Wt % CO_2	Approx. % crd in rock	Locality (lateral distance to opx isograd, if known)	Local conditions and/or facies	Mineralogy	Source*
177-1	-34.9	-41.8	6.9	1.15	40	Pikwitonei granulite domain, (PGD), N. Manitoba, Canada; N. Cauchon Lake (7.5 km)	648 °C, 6.1–6.8 kbar ²⁴ granulite	Qtz(rt), pl, kfs, bt, crd, gr, ilm, po, py, zrn	1
M218	-33.4	-40.0	6.6	0.91	30	PGD, N. Cauchon Lake, same outcrop as 177-1		Pl, kfs, bt, crd, gr, grt, rt, ilm	2
97-2	-11.2			1.90	50	PGD, N. Cauchon Lake (7–7.5 km)	774 °C, 6.1–6.8 kbar ²⁴ ; granulite	Qtz(rt), kfs, bt, crd, opx, Grt, po, zrn	1
97-7A	-14.9			1.56	25	PGD, N. Cauchon Lake, same outcrop as 97-2		Qtz, kfs, bt, crd, grt, py, ccp, zrn	1
231-4	-11.0			1.17	25	PGD, Prud'homme Lake (9 km)	750 °C, 6.1–6.8 kbar ²⁴ granulite	Crd, bt, rt, po, zen	1
M346	-31.8	-36.1	4.3	1.51	20	PGD, Prud'homme Lake (9 km)	780 °C, 6.4 kbar ²⁴ granulite	Qtz, kfs, bt, grt, crd, gr	2
M401	-32.9	-40.6	7.7	0.89	35	PGD, S. Cauchon Lake (0.5 km)	640 °C, 6.1–6.8 kbar ²⁴ ; granulite	Pl, kfs, bt, crd, gr, grt, sil	2
Bamble 3	-10.0			1.04	20	Bamble region, Norway	Amphibolite-granulite transition	Pl, bt, crd, grt, amphib	3
Grizzley creek WYO1	-12.7			0.88	50	Laramie Range, Wyoming	Amphibolite	Ky, crd, crn, (±sil, ged, mrg, st	4
	-14.6			0.37	75	Albany County, Wyoming	Granulite	Pl, kfs, bt, crd, opx, spl, mag, ilm, ap, zrn ²⁶	5, 6
135.489	-6.9			1.31		Manivitsy, Madagascar	Granulite ²⁷	Krn, bt, crd, sil, crn, spl, and, grd	7, 8
NE86A- 24A	-21.8	-25.1	3.3	1.28	20	Sturbridge, Massachusetts	675–730 °C, 6.3 kbar ²⁵	Qtz, pl, kfs, bt, crd, grt, Gr, sil, ru, zrn, opaques	9
NHQ-1	-21.7	-25.8	4.1	0.68	40	Nelson House, quarry Nelson Geras, Manitoba	Amphibolite	Qtz, kfs, bt, crd, gr, grt	1
Brazil 1	-7.4	-10.5	3.1	0.69		Minas Gerais, Governador Valadareao Brazil		Qtz, pl, bt, chl, crd, gr grt, amphib, ilm, rt, spl, s, ap, zrn	10
S. India 1	-20.8			0.92	30	Kerala, South India, Khondalite Belt	750 °C, 5–6 kbar ²⁸ Amphibolite-granulite transition	Pl, crd, grt, sil, spl	12
3083E	-18.5			0.32	75	Ellammankovilpatti, Tamil Nadu, India	750 °C, 5–6 kbar ²⁹ granulite	Krn, sil, hgb, crd, spl, mag, rt	7
3083D1	-20.5			0.32	75	Ellammankovilpatti, GECO mine,	750 °C, 5–6 kbar ²⁹ Amphibolite	Krn, sil, crd, rt, zrn, mnz	7
GECO	-15.1			0.40	85	Manitouwadge, Ontario		Crd, po, ccp, py, qtz, bt, felds	11

Mineral abbreviations: amphibole (amph), andalusite (and), apatite (ap), biotite (bt), chalcopyrite (ccp), chlorite (chl), cordierite (crd), corundum (crn), feldspar (felds), garnet (grt), gedrite (ged), grandierite (grd), graphite (gr), hobergite (hgb), ilmenite (ilm), k-feldspar (kfs), kornepine (krn), kyanite (ky), magnetite (mag), margarite (mrg), monazite (mnz), orthopyroxene (opx), plagioclase (pl), pyrite (py), pyrrhotite (po), quartz (qtz), rutile (rt), sillimanite (sil), spinel (spl), staurolite (st), sulfur (s), zircon (zrn).

* Sample sources: 1, J. Vry; 2, K. Mezger; 3, J. W. Valley; 4, R. Frost; 5, M. Cosca; 6, Cureton Mineral Co., Tucson, Az.; 7, E. Grew; 8, Musée Nationale d'Histoire Naturelle, Paris, France; 9, D. Moecher; 10, D. Perkins; 11, P. E. Brown; 12, T. Chacko.

Infiltration of a fluid equilibrated to mantle ($\delta^{13}\text{C} \approx -7$) or marble values ($\delta^{13}\text{C} = -7.2$ to $+7$)¹³ would homogenize any isotopic gradients and sharply increase $\delta^{13}\text{C}_{\text{crd}}$, even at extremely low CO_2/rock ratios, because most cordierite-bearing rocks have low total carbon. This fact is demonstrated in Fig. 1, which shows the minimum CO_2/rock ratio required to raise an initially low $\delta^{13}\text{C}_{\text{crd}}$ to $\delta^{13}\text{C}_{\text{crd}} = -15$ or above (solid lines), or to $\delta^{13}\text{C}_{\text{crd}} = -20$ or above (dashed lines). This figure was calculated using the following equation²⁰:

$$\frac{\text{CO}_2}{\text{rock}} = \left(\frac{\delta_{\text{crd}}^f - \delta_{\text{crd}}^i}{\delta_{\text{fluid}}^f - \delta_{\text{crd}}^f + \Delta} \right) \left(\frac{\text{weight 'equivalent CO}_2\text{'}}{\text{weight rock}} \right) \quad (1)$$

assuming an infiltrating fluid with initial $\delta^{13}\text{C}_{\text{fluid}} = -7$, final $\delta^{13}\text{C}_{\text{crd}} = -15$ or -20 , and $\Delta_{\text{crd-CO}_2} = 0$. (Superscripts i and f designate initial and final compositions, respectively.) Estimates of the approximate weight percentages of cordierite in the various samples were conservatively made, so that CO_2/rock ratios in Fig. 1 are upper limits. For samples containing graphite as well as cordierite, the graphite is expressed as an equivalent weight percentage of CO_2 .

Figure 1 shows that lower initial values of $\delta^{13}\text{C}_{\text{crd}}$ require larger amounts of infiltrating CO_2 to raise the $\delta^{13}\text{C}_{\text{crd}}$ to -15 or above. Again, these values are conservative, tending to overestimate CO_2/rock ratios. In most of the sample localities, as in the Pikwitonei domain, graphite is restricted to small, relatively isolated units. Such graphite should not seriously perturb the carbon isotope composition of a pervasively infiltrating CO_2 -rich fluid, for which CO_2/rock estimates are generally very large, ranging up to 0.08 (weight)²¹. Few rocks will have $\delta^{13}\text{C}_{\text{crd}} < -35$; however, even if the initial $\delta^{13}\text{C}_{\text{crd}}$ was ≤ -50 , the maximum predicted CO_2/rock would be < 0.06 for all the samples in Table

1 with final $\delta^{13}\text{C}_{\text{crd}} \leq -15$ (Fig. 1). For measured $\delta^{13}\text{C}_{\text{crd}} < -15$, or for infiltrating fluids with higher $\delta^{13}\text{C}$ values, the CO_2/rock would be even less. Hence, for most cordierites with $\delta^{13}\text{C}_{\text{crd}} < -15$ large amounts of pervasive carbonic fluid infiltration and exchange can be ruled out.

A calculation of CO_2/rock ratio for sample 'S. India 1', from the south Indian transitional granulite terrane, is illustrative of how very restrictive these calculations can be. If we set the original $\delta^{13}\text{C}$ of the rock at -25 and the cordierite content in the rock at 30%, then an infiltrating mantle fluid ($\delta^{13}\text{C} = -7$) would raise the rock $\delta^{13}\text{C}$ to values above the measured -20.8 if the CO_2/rock was ≥ 0.001 . If the infiltrating CO_2 was derived from marbles^{22,23}, then the required CO_2/rock ratios would be even smaller. It is important to recognize that the low CO_2/rock ratios estimated for this sample do not strictly rule out a process of local CO_2 infiltration into other rocks of the area. In fact, the sample locality is from the transitional amphibolite-granulite region which may have undergone 'patchy charnockitization'.

Carbon isotope analyses of 12 graphites from pelites and iron formations in the PGD give $\delta^{13}\text{C} = -24.7$ to -44.5 (mean -34.9). These distinctive values are comparable to other analyses of Archaean carbon from lower-grade graphitic shales, and are interpreted to represent original Archaean organic carbon values¹⁵. For graphite contents of < 0.2 wt %, and measured $\delta^{13}\text{C}_{\text{gr}} \leq -22$, Fig. 1 shows that CO_2/rock ratios must be ≤ 0.03 . Thus, any low- $\delta^{13}\text{C}$ graphite is also inconsistent with pervasive infiltration of large amounts of CO_2 .

In the PGD, values of $\delta^{13}\text{C}$ for cordierite channel CO_2 are unrelated to the mole fraction (J.K.V., unpublished data) or absolute content of CO_2 in cordierite, or to the lateral distance to the orthopyroxene isograd, but show a strong local depen-

dence on the presence or absence of graphite (-11 to -14.9 without graphite, -31.8 to -34.9 with graphite; see Table 1). Cordierites from other terranes also show a wide range in $\delta^{13}\text{C}$ (-21.7 to -6.9), again with little relationship to metamorphic grade. These data show that pelitic rocks from several granulite terranes have not been infiltrated by CO_2 in amounts greater than $\text{CO}_2/\text{rock} = 0.01$. This result indicates the importance of local fluid buffering in granulite formation, and is consistent with a model involving dehydration by partial melts⁸⁻¹⁰. Further detailed study will aid in determining the importance of limited amounts of localized CO_2 infiltration^{30,31}.

Received 24 September; accepted 29 December 1987.

1. Aines, R. D. & Rossman, G. *Am. Miner.* **69**, 319-327 (1984).
2. Armbruster, Th. & Bloss, F. D. *Am. Miner.* **67**, 284-291 (1982).
3. Armbruster, Th. & Bloss, F. D. *Nature* **286**, 140-141 (1980).
4. Armbruster, Th., Schreyer, W. & Hoefs, J. *Contr. Miner. Petrol.* **81**, 262-267 (1982).
5. Goldman, D. S., Rossman, G. R. & Dollase, W. A. *Am. Miner.* **62**, 1144-1157 (1977).
6. Hörmann, P. K., Raith, M., Raase, P., Ackermann, D. & Seifert, F. *Geol. Surv. Finland Bull.* **308**, 1-95 (1980).
7. Newton, R. C. in *Fluid-rock Interactions During Metamorphism* (eds Walther, J. V. & Wood, B. J.) 36-59 (Springer, New York, 1986).
8. Fyfe, W. S. *Phil. Trans. R. Soc. A273*, 457-461 (1973).
9. Valley, J. W. & O'Neil, J. R. *Contr. Miner. Petrol.* **85**, 158-173 (1984).
10. Lamb, W. M. & Valley, J. W. in *The Deep Proterozoic Crust in the North Atlantic Provinces* (eds Tobin, A. C. & Touret, J. L. R.) 119-131 (Reidel, Boston, 1985).
11. Valley, J. W. in *The Deep Proterozoic Crust in the North Atlantic Provinces* (eds Tobin, A. C. & Touret, J. L. R.) 217-236 (Reidel, Boston, 1985).
12. Johannes, W. & Schreyer, W. *Am. J. Sci.* **281**, 299-317 (1981).
13. Valley, J. W. in *Stable Isotopes in High Temperature Geological Processes* (eds Valley, J. W., Taylor, H. P. Jr & O'Neil, J. R.) 445-489 (BookCrafters, Chelsea, Michigan, 1986).

We thank all of the scientists who supplied samples (see Table 1). This study was funded by NSF, NASA and the Gas Research Institute.

Note added in proof: Two samples of crd-sill-grt-bt gneiss surrounded by charnockite were collected 6.1 km south of Kabbal, S. India with the aid of T. C. Devaraju, and yield $\delta^{13}\text{C}_{\text{crd}} = -11.8$, -11.0 , with 1.47 and 1.41 wt% CO_2 . These $\delta^{13}\text{C}_{\text{crd}}$ values are the highest yet measured in samples from S. India, demonstrating the local complexity that we predict, and possibly reflecting either an *in situ* sedimentary carbon source (mixed organics and carbonate) or externally introduced CO_2 .

14. Strauss, H. *Geochim. cosmochim. Acta* **50**, 2653-2662 (1986).
15. Morrison, J., Brown, P. E., Vry, J. K. & Valley, J. W. *Geol. Soc. Am. Abstr. Progm.* **18**, 700 (1986).
16. Valley, J. W. & O'Neil, J. R. *Geochim. cosmochim. Acta* **45**, 411-419 (1981).
17. Friedman, I. & O'Neil, J. R. *Prof. Pap. U.S. geol. Surv. No. 440-KK* (1977).
18. Schreyer, W. *Bull. mineral.* **108**, 273-291 (1985).
19. Kyser, T. K. in *Stable Isotopes in High Temperature Geological Processes* (eds Valley, J. W., Taylor, H. P. Jr & O'Neil, J. R.) 141-164 (BookCrafters, Chelsea, Michigan, 1986).
20. Taylor, H. P. Jr *J. geol. Soc. Lond.* **113**, 509-558 (1977).
21. Newton, R. C., Smith, J. V. & Windley, B. F. *Nature* **288**, 45-50 (1980).
22. Glassley, W. E. *Contr. Miner. Petrol.* **84**, 15-24 (1983).
23. Glassley, W. E. *Geochim. cosmochim. Acta* **47**, 597-616 (1983).
24. Mezger, K., Hanson, G. N. & Bohlen, S. R. *Contr. Miner. Petrol.* (in the press).
25. Robinson, P., Tracy, R. J., Hollocher, K. T., Schumacher, J. C. & Henry, N. B. IV *Univ. Mass. Amherst. Contrib.* **59**, 195-284 (1986).
26. Iiyama, T. *Bull. Soc. fr. Miner. Cristallogr.* **83**, 155-179 (1960).
27. Grew, E. S. *Z. geol. Wiss.* **14**, 525-558 (1986).
28. Chacko, T., Ravindra Kumar, G. R. & Newton, R. C. *J. Geol.* **95**, 343-358 (1987).
29. Grew, E. S., Abraham, K. & Medenbach, O. *Contr. Miner. Petrol.* **95**, 21-31 (1987).
30. Frost, B. R. & Frost, C. D. *Nature* **327**, 503-506 (1987).
31. Lamb, W. M. & Valley, J. W. *Nature* **312**, 56-58 (1984).

Uranium-series dating of the Mousterian occupation at Abric Romani, Spain

J. L. Bischoff*, R. Julia† & R. Mora‡

* U.S. Geological Survey, 345 Middlefield Road, Menlo Park, California 94025, USA

† Instituto de Geología Jaime Almera, Martí i Franqués, Barcelona 08028, Spain

‡ Museu d'Historia de la Ciutat, Carrer de la Força, Girona 17004, Spain

The precise evolutionary position of the Neanderthal people continues to be a major uncertainty in human evolution. Their origin and their relationship to anatomically modern people are unclear and are clouded by poor chronology. Lithic artefacts of the Mousterian type, found throughout Europe and the Mediterranean Basin, are believed to be the tool kit of the Neanderthals, but dates within Mousterian-bearing deposits are extremely rare. We report here on 20 high-quality uranium-series dates from Mousterian beds at Abric Romani, a rock shelter near Barcelona, Spain. The dates range from 39 to 60 kyr before present (BP) in an orderly stratigraphic succession and provide precise chronological control on an important Mousterian archaeological site.

The Mousterian industries were succeeded by Upper Paleolithic industries, the handiwork of anatomically modern people, between 35 and 40 kyr BP. This time range is at the very fringe of ^{14}C dating where the dates are particularly prone to error. Therefore, ^{14}C dates of the upper boundary of the Mousterian are usually considered minima. Mousterian is generally considered to extend at least back to the end of the last interglacial. But the chronology of cave sediments that represent this interglacial is uncertain between 75 and 125 kyr BP¹. There is evidence from La Chaise and Vaufray caves in France that the Mousterian might extend to before the last interglacial^{2,3}.

The method of choice for dating this time span is U-series, which is best applied to stalagmitic travertines (flowstones). Unfortunately, most of the classic Mousterian deposits lack travertine. Such travertines form primarily during the warm and

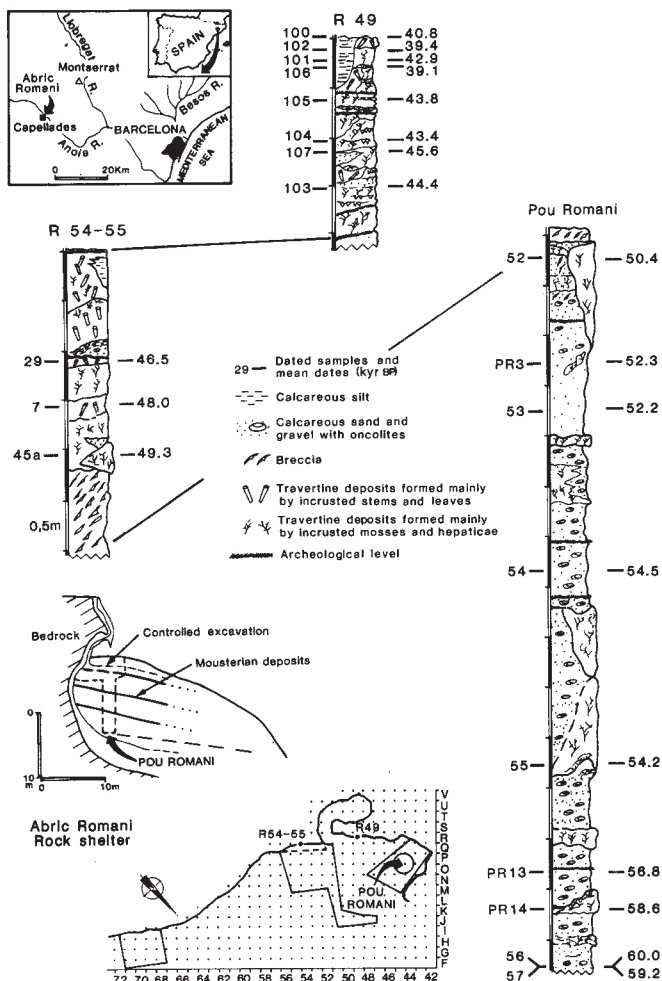


Fig. 1 Location, cross-section, excavation plan, and stratigraphic column of the rock shelter at Abric Romani. Sample numbers are shown on the left side of the stratigraphic column and corresponding U-series ages (in kyr) are shown on the right. Averaged ages are shown for those samples with multiple analyses.

Table 1 Uranium-series isotopic data and derived ages from samples from Abric Romani

Sample	Depth (cm)	U (p.p.m.)	$^{234}\text{U}/^{238}\text{U}$	$^{230}\text{Th}/^{234}\text{U}$	$^{230}\text{Th}/^{232}\text{Th}$	Age (kyr BP)	Corrected age* (kyr BP)
R-49 section							
100	+30	1.5	2.23 ± 0.05	0.35 ± 0.01	10.8	44.9 ± 1.5	40.0
		1.01	2.33 ± 0.03	0.34 ± 0.005	74.2	44.4 ± 0.7	42.8
		0.68	2.39 ± 0.05	0.34 ± 0.01	12.6	43.1 ± 1.5	39.2
		0.75	2.30 ± 0.05	0.35 ± 0.01	9.1	45.6 ± 1.5	41.4
102	-10	1.24	2.28 ± 0.05	0.36 ± 0.01	7.4	46.8 ± 1.5	39.4
101	-20	1.3	2.18 ± 0.05	0.37 ± 0.01	10.8	48.2 ± 1.6	42.9
106	-30	0.76	2.17 ± 0.04	0.34 ± 0.01	8.5	43.5 ± 1.5	39.1
105	-52	1.99	2.30 ± 0.04	0.36 ± 0.01	24.8	46.2 ± 1.5	43.8
104	-102	0.73	2.33 ± 0.04	0.36 ± 0.01	8.7	46.0 ± 1.5	43.4
107	-120	0.88	2.16 ± 0.04	0.39 ± 0.01	20.9	50.9 ± 1.6	48.1
		0.60	2.25 ± 0.05	0.35 ± 0.01	19.2	45.1 ± 1.5	43.1
		0.69	2.22 ± 0.05	0.35 ± 0.01	52.4	44.2 ± 1.5	—
103	-123	0.33	2.30 ± 0.07	0.35 ± 0.01	>1,000	44.6 ± 1.5	—
R 54-55 section							
29	-275	0.92	2.30 ± 0.05	0.35 ± 0.01	70.4	45.3 ± 1.5	—
		0.79	2.15 ± 0.04	0.37 ± 0.01	105.2	47.7 ± 1.6	—
7	-307	0.96	2.19 ± 0.04	0.37 ± 0.01	86.2	48.0 ± 1.6	—
45a	-375	0.55	2.27 ± 0.05	0.38 ± 0.01	>1,000	49.3 ± 1.6	—
Pou Romani section							
52	-520	1.82	2.24 ± 0.03	0.38 ± 0.01	>1,000	50.0 ± 1.6	—
		1.80	2.27 ± 0.03	0.39 ± 0.005	>1,000	50.8 ± 0.8	—
PR-3	-650	2.00	2.26 ± 0.04	0.40 ± 0.01	>1,000	52.0 ± 1.26	—
		1.84	2.24 ± 0.02	0.40 ± 0.005	>1,000	53.0 ± 0.8	—
		1.00	2.23 ± 0.03	0.39 ± 0.005	>1,000	51.9 ± 1.6	—
53	-685	0.65	2.25 ± 0.06	0.40 ± 0.01	>1,000	52.2 ± 1.6	—
54	-820	1.34	2.19 ± 0.04	0.41 ± 0.01	27.5	54.9 ± 1.7	—
		1.22	2.35 ± 0.05	0.41 ± 0.01	26.7	54.1 ± 1.6	—
55	-1,010	1.56	2.10 ± 0.04	0.43 ± 0.015	13.0	58.5 ± 2.6	55.0
		0.80	2.20 ± 0.04	0.40 ± 0.01	30.0	53.4 ± 1.6	—
		1.60	2.13 ± 0.04	0.45 ± 0.01	31.9	60.6 ± 1.7	—
PR-13	-1,155	1.46	2.20 ± 0.03	0.42 ± 0.01	26.1	55.5 ± 1.7	—
		0.92	2.18 ± 0.04	0.41 ± 0.01	29.5	54.5 ± 1.7	—
		1.93	2.19 ± 0.04	0.43 ± 0.005	39.8	57.2 ± 0.8	—
PR-14	-1,180	1.79	2.16 ± 0.04	0.44 ± 0.01	29.4	59.6 ± 1.7	—
		0.97	2.15 ± 0.03	0.44 ± 0.01	27.5	59.0 ± 1.7	—
		1.73	2.13 ± 0.02	0.46 ± 0.005	22.9	63.2 ± 0.9	61.0
56	-1,240	0.91	2.23 ± 0.04	0.44 ± 0.01	42.7	59.0 ± 1.7	—
		1.75	2.24 ± 0.05	0.46 ± 0.01	14.3	63.3 ± 1.8	60.1
57	-1,240	0.93	2.22 ± 0.05	0.45 ± 0.015	14.9	61.3 ± 2.6	58.0
		0.89	2.31 ± 0.08	0.46 ± 0.015	16.3	62.5 ± 2.6	59.6
Carbonate-mud from living moss (Capellades town park)							
		2.51	2.66 ± 0.07	0.02 ± 0.005	0.9	0.24 ± 0.6	<2.4
Limestone basement (La Bofia)							
		0.58	1.26 ± 0.04	1.13 ± 0.04	9.0	>350	—
		0.58	1.25 ± 0.04	1.10 ± 0.04	19.8	>350	—
Insoluble residue (sample 106)							
		2.69	1.12 ± 0.08	1.12 ± 0.08	1.01	>350	—

Error ($\pm$) based on 1σ uncertainty from counting statistics.

* Ages corrected by method of Ku and Liang¹¹ scheme II, using isotopic compositions of insoluble residue of sample 106.

humid interglacial times, and are generally lacking within the glacial climate of the Mousterian levels^{4,5}. Therefore, dates from within Mousterian beds are rare. One exception is thermoluminescence dates of burned flints from the type deposit at Le Moustier, France with ages spanning 40.3 to 55.8 kyr BP⁶.

Abric Romani, ~45 km northwest of the city of Barcelona, is situated on a 90-m limestone cliff which separates the town of Capellades from the river Anoia (Fig. 1). The exposed part of the deposit consists of over 13 m of carbonate sediments containing Mousterian industries. Originally the section was overlain by 3 m of red silt-rich beds containing Upper Paleolithic industries, but these have all been removed. The 13 m of Mousterian beds are largely intact.

The deposit was originally discovered by Amador Romani in 1909, who carried out excavations from time to time over the

succeeding 20 years and exhausted the Upper Paleolithic layers. In 1956 E. Ripoll, H. de Lumley and G. Laplace reinitiated excavation that continued until 1961. This later excavation extended into the first few meters of the Mousterian layers, the industry of which was attributed to Denticulate Mousterian⁷. Upper Paleolithic artefacts were attributed to Gravettian and Aurignacian⁸.

Excavations have been carried out since 1983 by E. Carbonell, R. Mora and A. Cebria (Centro de Recerques Paleo-Eco-Socials, Museu d'Historia de la Ciutat, Girona, Spain) who rediscovered a test shaft (Pou Romani) made earlier by Amador Romani that exposes an additional 10 m of Mousterian-bearing beds. Fifteen archaeological levels, each rich in bone, tools and charcoal, are recognized throughout the exposed section, and these represent sporadic short-duration occupations, separated by long periods

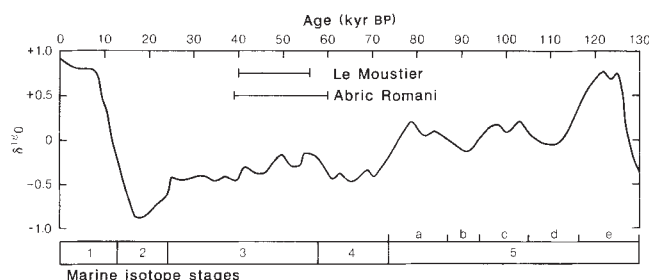


Fig. 2 The time span of the dated Mousterian layers from Abric Romani and Le Moustier are shown with respect to variation of $\delta^{18}\text{O}$ (normalized) of the Marine Isotope Stages 1–5 (ref. 13).

of carbonate sedimentation. Artefacts were the product of centripetal flaking, and are dominated by flakes and flake debris, with fewer shaped or retouched tools. Among the retouched tools are denticulates, side-scrapers and notched tools. There are also many undifferentiated flakes with retouching which cannot be ascribed to any well defined type. No bifaces have been found. The fauna include *Equus caballus*, *Cervus elphus*, *Capra pyrenaica*, *Equus hydruntinus*, *Bos sp.*, *Oryctolagus cuniculus*, *Canis lupus*, *Hyaena spelaea* and *Felis sylvestris*. Excavation has not yet reached the base of the Mousterian levels, and several more metres of section may exist beneath the exposed section in the test shaft.

The deposit is primarily composed of well-bedded carbonate sediments that formed when stream run-off extended over the face of the cliff from the plateau above, wetting the roof and back walls as well as the floor of the rock shelter. Carbonates precipitated on all of these surfaces by the photosynthetic removal of CO_2 from the water by aquatic plants and mosses. Fine structure of the moss and imprints of other plant debris, including leaves and aquatic grasses, are well preserved in the carbonates⁹. The beds are porous, friable, and poorly crystalline, and are quite distinct from the highly-crystalline stalagmitic travertines normally used for dating in cave environments. Twenty samples for dating were taken spanning 13 m of stratigraphic section (Fig. 1). Thirty two analyses were performed, including replicates, by conventional techniques using α -spectroscopy¹⁰.

Isotopic composition of the samples and the stratigraphic succession of ages indicate a high validity for the dates. Ages increase evenly from 39 kyr at the top to 60 kyr at the bottom of the section (Fig. 1, Table 1). Uranium content is relatively high ranging between 1 and 2 p.p.m., and the $^{234}\text{U}/^{238}\text{U}$ ratio is remarkably uniform between 2.1 and 2.3. Detrital contamination of Th is negligible based on the high $^{230}\text{Th}/^{232}\text{Th}$ ratios and insoluble residues of less than 1%.

The ages of those examples with $^{230}\text{Th}/^{232}\text{Th}$ ratios less than 20, primarily those from the uppermost 1.5 m of section, were corrected for slight detrital contamination¹¹ using the isotopic composition of the insoluble residue from sample 106. The isotopic composition of this residue is essentially at equilibrium rendering the correction rather straightforward and making the ages only 2–7 kyr younger (Table 1). Multiple analyses of several samples show agreement within the range of the counting error (Table 1), which yields a typical uncertainty of about ± 2 kyr. A ^{231}Pa age of 55.4 ± 6 kyr was obtained for the bottom of the section (sample 56) by analysis of ^{227}Th ¹². This date is within the counting error of the ^{230}Th age of 61.0 ± 2.5 kyr for the same sample (Table 1). Therefore, the age is confirmed by the internal concordance between two independent decay chains. Also analysis of the limestone bedrock shows full isotopic equilibrium, and precipitates from a nearby modern moss yields a modern age (Table 1). All indications are that the system at Abric Romani is fully closed with respect to uranium-series isotopes and that the dates are reliable. The reasons for successful uranium-series dating are probably a high uranium content in the waters and a high efficiency of uranium fixation caused

by the high organic content of the sediments.

The exposed section, therefore, represents 20,000 years of fairly constant climatic conditions and a continuous and rapid sedimentation rate (60 cm kyr^{-1}). The dating shows no abrupt changes in sedimentation rate, and wet conditions, at least locally, appear to characterize the entire time span. A trend to dryer conditions toward the top of the section, however, is suggested by the higher percentage of insoluble residues (eolian silt and clay) of the uppermost 1.5 m. A rapid change to much dryer conditions after this time is indicated by the succession to silt-rich beds with Upper Paleolithic industries immediately after 39 kyr BP.

The time span of the Mousterian at Abric Romani encompasses that of Le Moustier, and both time spans are shown with respect to the marine isotope stage chronology in Fig. 2. The beds at Le Moustier fall within isotope stage 3, whereas those from Abric Romani extend from the middle of stage 3 back into stage 4. The section at Le Moustier lacks carbonates and the sedimentology indicates deposition occurred during a dry and cold period^{4,6}. Abric Romani appears to represent wetter conditions, suggesting the climate was considerably different south and north of the Pyrenees during isotope stage 3. This is not surprising considering that the climates of the two regions today are also dissimilar.

The dating at Abric Romani establishes the Mousterian back to at least 60 kyr BP in this region of Spain. Whether it extends to isotope stage 5e (125 kyr) or earlier may be answered in part by excavations now being carried out into the deeper parts of the section at Abric Romani.

We thank L. Freeman, J. Gardner, A. Jelinek, J.-P. Rigaud, L. Straus and B. Szabo for helpful comments.

Received 22 October 1987; accepted 11 January 1988.

1. Dennell, R. *Nature* **301**, 199 (1983).
2. Blackwell, B., Schwarcz, H. P. & Debenath, A. *J. archaeol. Sci.* **10**, 493–513 (1983).
3. Blackwell, B. & Schwarcz, H. P. in *Grotte Vaufray: Paléoenvironnements, Chronologie et activités Humaines* (ed. Rigaud, J.-Ph) (Société Préhistorique Française, Paris) (in the press).
4. Laville, H., Rigaud, J.-P. & Sackett, J. *Rock Shelters of the Perigord: Geological Stratigraphy and Archaeological Succession* (Academic, New York, 1980).
5. Harmon, R. S., Ford, D. C. & Schwarcz, H. P. *Can. J. Earth Sci.* **14**, 2543–2552 (1977).
6. Valladas, H., Geneste, J. M., Joron, J. L. & Chadelle, J. P. *Nature* **322**, 452–454 (1986).
7. de Lumley, H. & Ripoll, E. *L'Anthropologie*, **T. 66**, (1–2) 1–35 (1962).
8. Laplace, G. *L'Anthropologie*, **T. 66**, (1–2), 36–43 (1962).
9. Julia, R. in *Carbonate Depositional Environments* (eds Scholle, P., Bebout, D. & Moore, C.) *Memoir 33*, 64–72 (Am. Ass. Petrol. Geologists, Tulsa, 1983).
10. Ku, T. L. thesis, Univ. Columbia (1986).
11. Ku, T. L. & Liang, Z.-C. *Nucl. Instrum. Meth.* **563**–571 (1984).
12. Moore, W. S. & Somayajulu, B. L. K. *J. geophys. Res.* **79**, 5065–5068 (1974).
13. Martinson, D. G. *et al. Quat. Res.* **27**, 1–24 (1987).

A response by benthic Foraminifera to the deposition of phytodetritus in the deep sea

Andrew J. Gooday

Institute of Oceanographic Sciences, Deacon Laboratory (NERC), Wormley, Godalming, Surrey GU8 5UB, UK

A recent major discovery has been the rapid sedimentation of phytodetritus to the deep-sea floor^{1–3}. Although benthic mega-faunal invertebrates appear to seek out this relatively fresh food source^{1,4}, and its seasonal arrival on the sea floor may synchronize reproduction in some echinoderms⁵, a convincing response by benthic organisms to phytodetritus has not been demonstrated³. Here I present evidence that certain small benthic Foraminifera (within the meiofaunal size-range) react dramatically to the presence of phytodetritus. Fresh aggregates of this material harbour abundant, low-diversity populations of these protists. The three commonest species are usually poorly represented in the more diverse assem-

Table 1 Gross taxonomic composition of the benthic foraminiferal assemblages

Meteor station	172		179		200	
Depth	4,483 m		4,538 m		4,523 m	
Date	7 August 1986		8 August 1986		14 August 1986	
	Core 1		Core 2			
Sample	Surface sediment	Phyto-detritus	Surface sediment	Phyto-detritus	Surface sediment	Phyto-detritus
Allogromiidae	24.8	45.4	18.8	44.7	14.8	27.9
Saccamminidae (soft-walled) ³⁰	20.8	1.1	23.1	—	20.1	—
Saccamminidae (others)	18.4	—	8.7	—	10.7	—
Astrorhizacea (tubes)	15.2	—	7.3	0.3	14.1	1.1
Komokiacea	2.4	—	0.9	—	2.0	—
<i>Globigerina</i> inhabitants*	2.4	—	0.7	—	4.7	—
Hormosinacea	4.0	—	5.0	0.3	4.7	—
Other multichambered Textulariina	6.4	13.8	22.7	9.3	12.8	11.5
Miliolina	—	2.2	0.7	0.3	2.0	8.7
Lagenina	3.2	0.4	—	—	1.7	—
Rotaliina	3.2	37.0	12.1	45.0	6.0	50.8
Total Foraminifera†	125	275	437	313	149	183
Foraminiferal species†	48	12	57	11	52	11
Total metazoan meiofauna†	37	21	148	34	69	20
Sample volume (ml)	0.9	0.5	2.5	0.7	1.2	0.5

The overlying water was carefully drained to within ~2 cm of the sediment surface. The phytodetritus aggregates were then picked off the surface with soft forceps and shaken while in water to dislodge any adhering sediment. The surface sediment (top 1–2 mm free from phytodetritus) was removed with a pipette. All of the samples were preserved in buffered 4% formaldehyde. The sediment was processed and sorted as in ref. 31. Phytodetritus aggregates were placed in water in a 15 mm-diameter cavity slide and gently teased apart to reveal the inhabiting organisms. In view of problems involved with the use of Rose Bengal^{28,31}, all of the specimens with red-stained contents were examined under a high-power light microscope in glycerol (which renders the test more transparent) to ensure that the stained material was fresh foraminiferal protoplasm. Figures are percentages unless stated otherwise.

* Benthic Foraminifera which are obligate inhabitants of empty *Globigerina* shells³¹.

† Number per sample.

Table 2 The percentage abundance of all phytodetritus-inhabiting benthic Foraminifera in sediment and phytodetritus samples

Meteor station	172		179		200	
	Core 1		Core 2			
Sample	Surface sediment	Phyto-detritus	Surface sediment	Phyto-detritus	Surface sediment	Phyto-detritus
Allogromiid sp.*	0.8	45.5	2.7	45.0	—	7.4
Soft-walled saccamminid 1	8.8	0.4	6.6	—	0.7	2.5
Soft-walled saccamminid 2	0.8	0.4	—	—	—	—
Saccamminid sp.	—	—	—	—	—	9.9
<i>Hyperammina</i> sp.	—	—	4.8	0.3	5.4	0.5
<i>Spiroplectammina biformis</i>	—	0.7	1.1	—	—	1.2
<i>Trochammina</i> sp.	2.4	5.1	11.7	9.3	6.0	8.7
<i>Adercotryma glomerata</i>	2.4	8.0	5.0	—	3.4	2.7
<i>Pyrgo serrata</i>	—	—	—	—	—	0.5
<i>Pyrgoella</i> sp.	—	2.2	0.2	0.3	—	8.2
<i>Fissurina</i> sp.	0.8	0.4	—	—	0.7	—
<i>Bulimina</i> sp.	0.8	1.1	0.5	2.2	0.7	—
<i>Epistominella exigua</i> *	—	13.8	5.5	19.5	1.3	13.1
<i>Eponides weddellensis</i> *	0.8	21.8	1.8	18.5	—	35.0
<i>Globocassidulina subglobosa</i>	0.8	0.9	0.7	2.9	0.7	1.1
<i>Planulina wuellerstorfi</i>	—	—	—	0.6	—	—
? <i>Nonionella</i> sp.	—	—	—	1.3	—	0.5

Columns headed surface sediment only include those species which occur in both sediment and phytodetritus samples; hence the percentages in these columns total <100%. Consistently abundant species are indicated by asterisk.

blages inhabiting the underlying sediment. These findings suggest that some deep-sea benthic Foraminifera, like their shallow-water relatives^{6–8}, are specialist feeders that bloom opportunistically when the appropriate food (phytodetritus and associated micro-organisms) becomes available, while others remain unaffected by the organic influx.

One aim of the BIOTRANS Programme of the Institut für Hydrobiologie und Fischereiwissenschaft, Hamburg, is to investigate organic fluxes in an abyssal area of the northeast Atlantic (47°00'–47°30' N, 19°00'–20°00' W; 3,800–4,600 m bathymetric depth). During the BIOTRANS IV cruise (July–

August 1986) phytodetritus was observed on cores obtained with a multiple corer⁹, as a partial or complete surface layer or as isolated aggregates. The phytodetritus was macroscopically similar to material described earlier from the Porcupine Seabight^{1–3}, although less degraded. Full details of its composition will be published elsewhere (H. Thiel *et al.*, manuscript in preparation).

The results reported here are based on samples of phytodetritus and superficial sediment (upper 1–2 mm) taken from cores recovered at three stations on the abyssal plain. The four sediment samples examined contained 125–437 live (Rose-

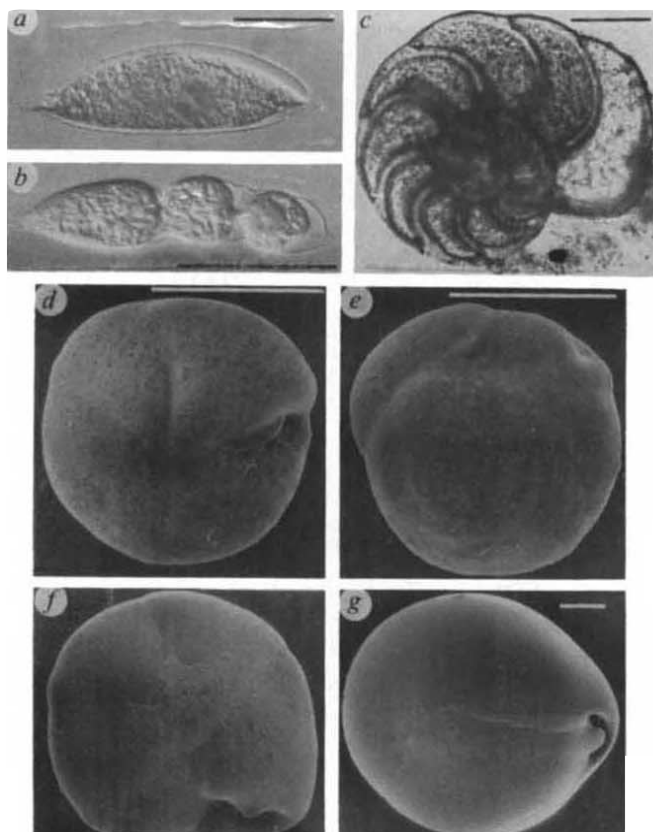


Fig. 1 Phytodetritus-inhabiting benthic Foraminifera. *a-c*, Specimens mounted in glycerol and photographed by transmitted light; *d-g*, scanning electron microscope (SEM) micrographs. *a*, Allogromiid; undescribed species with thin, transparent organic test wall, large nucleus, and two opposing apertures located at ends of short tubes. *b*, Specimen with constricted test and protoplasm. Constrictions are evident in 13% of specimens and appear not to be artefacts, similar structures having been described in near-shore allogromiids³². Presence of single nucleus suggests constrictions are due to morphological plasticity³² and are not preliminary stages in reproductive fission. *c*, *Planulina wuellerstorfi*. *d, e*, *Eponides weddellensis*. *f*, *Epistominella exigua*. *g*, *Pyrgoella* sp. Named species were identified by reference to type specimens in British Museum (Natural History) London, and Royal Museum of Scotland, Edinburgh. Scale bars = (*a, b*) 50 μ m; (*c-g*) 100 μ m.

Bengal stained) Foraminifera (69–77% of total meiofauna) (Table 1). Between 34 and 57 putative species were recognized and the Shannon–Weaver diversity index ($H(S)$) ranged from 2.83 to 3.62. Important higher taxa were the suborders Rotaliina and Textulariina, the superfamily Astrohizacea and the families Allogromiidae and Saccamminidae (Table 1). Phytodetritus samples also yielded a rich collection of living (stained) benthic Foraminifera (88–93% of all colonizing organisms). Planktonic Foraminifera containing protoplasm were sometimes encountered within the aggregates but were distinguished readily from benthic species. The phytodetritus assemblages were dominated by the Allogromiidae and Rotaliina, both taxa being considerably more abundant than in the sediment (Table 1). But species diversity was much lower (11–12 species per sample, $H(S)$ = 1.50–1.69). Also, the three consistently most abundant species, *Epistominella exigua* (Brady), *Eponides weddellensis* Earland (Rotaliina) and an undescribed allogromiid (Fig. 1), occurred only sporadically in the underlying sediment (Table 2). Two less abundant species, *Planulina wuellerstorfi* (Schwager) (Rotaliina) and *Pyrgoella* sp. (Miliolina), were found mainly in the phytodetritus. Only *Adercotryma glomerata* (Brady) and *Trochammina* sp. (multichambered Textulariina) were fairly common on both sediment and phytodetritus samples.

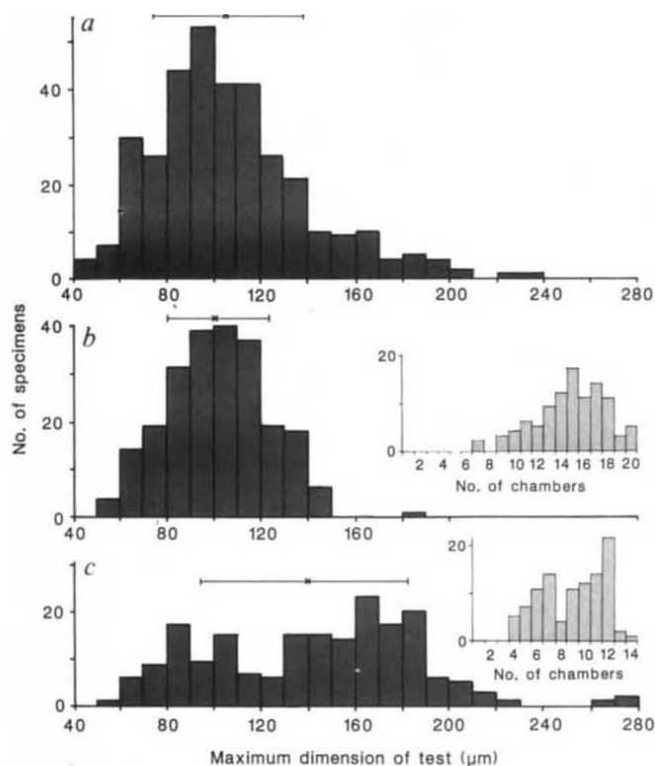


Fig. 2 Size distributions of living benthic Foraminifera from the phytodetritus. Means and standard deviations are given above each histogram. *a*, Allogromiid (339 observations). *b*, *Eponides weddellensis* (228 observations). *c*, *Epistominella exigua* (185 observations). Insets show distribution of numbers of chambers per test for *Eponides weddellensis* (102 observations) and *E. exigua* (102 observations) (allogromiid is unilocular). Specimens were mounted in glycerol, and maximum dimensions measured to nearest 2.5 μ m under compound microscope.

The phytodetritus probably provides a living space free, initially, from potential competitors and also a food source for the Foraminifera. Calcareous species (including *E. exigua* and *E. weddellensis*) have greenish protoplasm suggesting that they ingest phytodetritus. Also, a preliminary study using epifluorescence microscopy (C. M. Turley, personal communication) demonstrates that all three dominant species consume prokaryotic microorganisms which occur in large numbers on the phytodetritus (Lochte and Turley, manuscript submitted). This important finding is being investigated further (Gooday and Turley, manuscript in preparation). The Foraminifera must be consumed, in turn, by larger organisms such as scaphopod molluscs¹⁰, and some holothurians which selectively ingest phytodetritus-rich surface sediments⁴. Hence, they may provide a vital trophic link between prokaryotes and metazoans, perhaps playing a role analogous to that of planktonic microflagellates¹¹. This would be consistent with the known ecological importance of protists in many other environments¹².

The utilization of phytodetritus by deep-sea benthic Foraminifera is consistent with observations on shallow-water species. These protists are abundant in high-productivity areas¹³. Unusually large, low-diversity populations occur near nutrient sources such as sewage outfalls¹⁴ or on patches of decaying algae⁶. Reproduction and population growth sometimes follow phytoplankton blooms^{15–17}. Some salt-marsh species are opportunists that flourish under favourable conditions (for example, when their preferred food source becomes abundant and competitors are rare), whereas others are more general feeders, maintaining fairly stable populations in a fluctuating environment^{6–8,17}. My findings suggest that this kind of niche specialization also exists among deep-sea Foraminifera.

These large populations may have resulted from (1) the migration of Foraminifera into the phytodetritus from the sediment, (2) reproduction inside the aggregates, or (3) a combination of these mechanisms. The size distributions of the dominant species shed some light on this question. The allogromiid has a rather broad (40–240 μm), unimodal size spectrum, dominated by small (<140 μm), possibly juvenile individuals; in *E. exigua* the test size and the number of constituent chambers both show a bimodal distribution, presumably reflecting the presence of two generations (Figs 2a, c). Although such patterns are difficult to interpret without comparable data from other seasons, the abundance of small specimens does suggest active reproduction in these species^{15,18}. *E. weddellensis*, on the other hand, has a fairly symmetrical, unimodal size curve (Fig. 2b).

These results provide the first evidence that deep-sea meiofaunal organisms can respond to low-intensity ecological disturbances of the kind caused by phytodetritus deposition. The 1986 spring bloom occurred during May in the BIOTRANS area (Continuous Plankton Recorder data, J. M. Colebrook, Institute for Marine Environmental Research, personal communication). Phytodetritus originating from this bloom would have reached the sea floor at 4,500 m during June or early July, assuming a sinking rate of 100–150 m per day². The foraminiferal response to it probably occurred, therefore, within a period of 1–2 months. This rapid reaction supports recent evidence^{19,20} that deep-sea communities can respond to disturbances more quickly than previously believed. Natural disturbances such as nekton falls¹⁹, the formation of biogenous mounds²⁰, and strong near-bottom currents²¹ are thought to influence the structure of deep-sea communities and help maintain their high diversity. The phytodetritus assemblages have a low diversity and are dominated by species which are uncommon in the underlying sediment. Deep-sea macrofaunal assemblages living in disturbed sediment display some similar features^{19,21–23}.

In addition to their biological significance, these data may have relevance in Caenozoic palaeo-oceanography. Correlations between the abundances of certain benthic foraminiferal species, and a high organic input to the sediment, have been recognized under both modern and Pleistocene upwelling areas^{24–27}. In non-upwelling areas, phytodetritus appears to trigger a similar population increase among opportunistic species able to utilize it. The observation that these species live within phytodetritus aggregates augments the meagre existing information on the ecology of deep-sea benthic Foraminifera. This is important because microhabitat preferences seem to influence the stable isotope (¹⁸O and ¹³C) content of calcareous tests^{27–29}, an increasingly valuable tool in palaeo-oceanography.

I thank H. Thiel for inviting me to participate in BIOTRANS IV, M. V. Angel, D. S. M. Billett, A. L. Rice, H. Thiel, A. Thistle, D. Thistle, M. H. Thurston and C. M. Turley for reviewing the manuscript. This is the fifth BIOTRANS publication.

Received 30 November 1987; accepted 28 January 1988.

1. Billett, D. S. M., Lampitt, R. S., Rice, A. L. & Mantoura, R. F. C. *Nature* **302**, 520–522 (1983).
2. Lampitt, R. S. *Deep Sea Res.* **32**, 885–897 (1985).
3. Rice, A. L. *et al. Proc. R. Soc., Edinb.* **88B**, 265–279 (1986).
4. Billett, D. S. M., Llewellyn, C. & Watson, J. in *Proceedings of the 6th International Echinoderm Conference* (ed. Burke, R. D.) (Balken, Rotterdam, in the press).
5. Tyler, P. A. *Proc. R. Soc., Edinb.* **88B**, 175–190 (1986).
6. Lee, J. J., Muller, W. A., Stone, R. J., McEnery, M. E. & Zucker, W. *Mar. Biol.* **4**, 44–61 (1969).
7. Lee, J. J. in *Foraminifera*, Vol. 2 (eds Hedley, R. H. & Adams, C. G.) 207–260 (Academic, New York, 1974).
8. Muller, W. A. *Mar. Biol.* **31**, 339–351 (1975).
9. Barnett, P. R. O., Watson, J. & Connelly, D. *Oceanologica Acta* **7**, 399–408 (1984).
10. Morton, J. E. *J. mar. biol. Ass. U.K.* **38**, 225–238 (1959).
11. Johnson, P. W., Huai-Shu, X. & Sieburth, J. McN. *Ann. Inst. Oceanogr., Paris* **58**(S), 297–308 (1982).
12. Fenchel, T. *Ecology of Protozoa* (Science Tech, Madison; Springer, Berlin, 1987).
13. Phleger, F. B. *Marit. Sediments Spec. Pub.* **1**, 107–117 (1976).
14. Bandy, O. L., Ingle, J. C. & Resig, J. M. *Limnol. Oceanol.* **9**, 112–123 (1964).
15. Wefer, G. *Marit. Sediments Spec. Pub.* **1**, 39–50 (1976).
16. Faubel, A., Hartwig, E. & Thiel, H. *'Meteor' Forsch.-Ergebn.* **D36**, 35–48 (1983).
17. Lipps, J. H. in *Biotic Interactions in Recent and Fossil Benthic Communities* (eds Tevesz, M. J. S. & McCall, P. S.) 331–376 (Plenum, New York, 1983).
18. Murray, J. W. *J. foram. Res.* **13**, 1–12 (1983).
19. Smith, C. R. *J. Mar. Res.* **44**, 567–600 (1986).

20. Smith, C. R., Jumars, P. A. & DeMaster, D. J. *Nature* **323**, 251–253 (1986).
21. Thistle, D., Yingst, J. Y. & Fauchald, K. *Mar. Geol.* **66**, 91–112 (1985).
22. Grassle, J. F. *Nature* **265**, 618–619 (1977).
23. Desbruyeres, D., Bervas, J. Y. & Khripounoff, A. *Oceanol. Acta* **3**, 285–291 (1980).
24. Sen Gupta, B. K., Lee, R. F. & May, M. S. *J. Paleont.* **55**, 853–857 (1981).
25. Caralp, M.-H. *Oceanol. Acta* **7**, 509–515 (1984).
26. Lutze, G. F., Pflaumann, U. & Weinholz, P. *'Meteor' Forsch.-Ergebn.* **C40**, 163–180 (1986).
27. Zahn, R., Winn, K. & Sarnthein, M. *Paleoceanography* **1**, 27–42 (1986).
28. Corliss, B. H. *Nature* **314**, 435–438 (1985).
29. Grossman, E. L. *J. foram. Res.* **17**, 48–61 (1987).
30. Gooday, A. J. *Sarsia* **71**, 275–287 (1986).
31. Gooday, A. J. *Deep Sea Res.* **33**, 1345–1373 (1986).
32. Nyholm, K.-G. *Zool. Bidr. Upps.* **30**, 465–474 (1955).

Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus

Andrew J. King, Mary E. Hutchings, David R. Moore & Colin Blakemore

University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK

Environmental factors play an important role in certain aspects of the development of sensory systems^{1–3}. But the way in which the maturation of different sensory modalities is coordinated is poorly understood. We have investigated this question neurophysiologically in the mammalian superior colliculus (SC), which contains topographically aligned maps of visual and auditory space^{4–6}. We report here that an essentially normal auditory map, in approximate register with the visual map, is found in the SC of adult ferrets reared with abnormal binaural localization cues. Also, if, early in life, one eye is deviated laterally, there is a compensatory shift in the auditory map, but early eye rotation totally disorders the auditory representation. These results imply that development of the auditory map is affected by visual activity or by information about eye position and that there is definite, but limited, capacity for the auditory map to reorganize so that it remains aligned with the visual map.

Recent behavioural experiments have shown that young barn owls can adjust their sound localization to overcome the distorting effects of plugging one ear; they can also readjust their orienting reactions after plug removal, but only if normal visual experience is available^{7–9}. Monaural occlusion during development has also been shown to cause changes in the properties of auditory neurons in the owl's optic tectum, so as to preserve a normal map of auditory space^{10,11}. But the neural basis of the influence of vision on sound localization has not been examined. We have now investigated whether functional interaction between different sensory systems during development plays a part in establishing or maintaining the registration of visual and auditory maps in the mammalian superior colliculus.

All forms of sensory perturbation (see Figs 1–3) were started on postnatal days 23–28, several days before the functional onset of vision¹² and hearing¹³ in the ferret. When the animals were 3–16 months old, they were anaesthetized with 18 mg kg⁻¹ intramuscular (i.m.) alphaxalone/alphadolone acetate, prepared for extracellular recording and paralysed with 12 mg intravenous (i.v.) gallamine triethiodide, as previously described⁶. Anaesthesia and paralysis were maintained by i.v. infusion of 1 mg kg⁻¹ h⁻¹ sodium pentobarbital and 20 mg kg⁻¹ h⁻¹ gallamine triethiodide. The pinnae were adjusted to the normal position before recording, on the basis of measurements made before surgery. No changes in eye position (within $\pm 2^\circ$) were detected following paralysis.

Acoustically-responsive neurons in the deep layers of the SC of normal adult ferrets respond best to sounds from particular

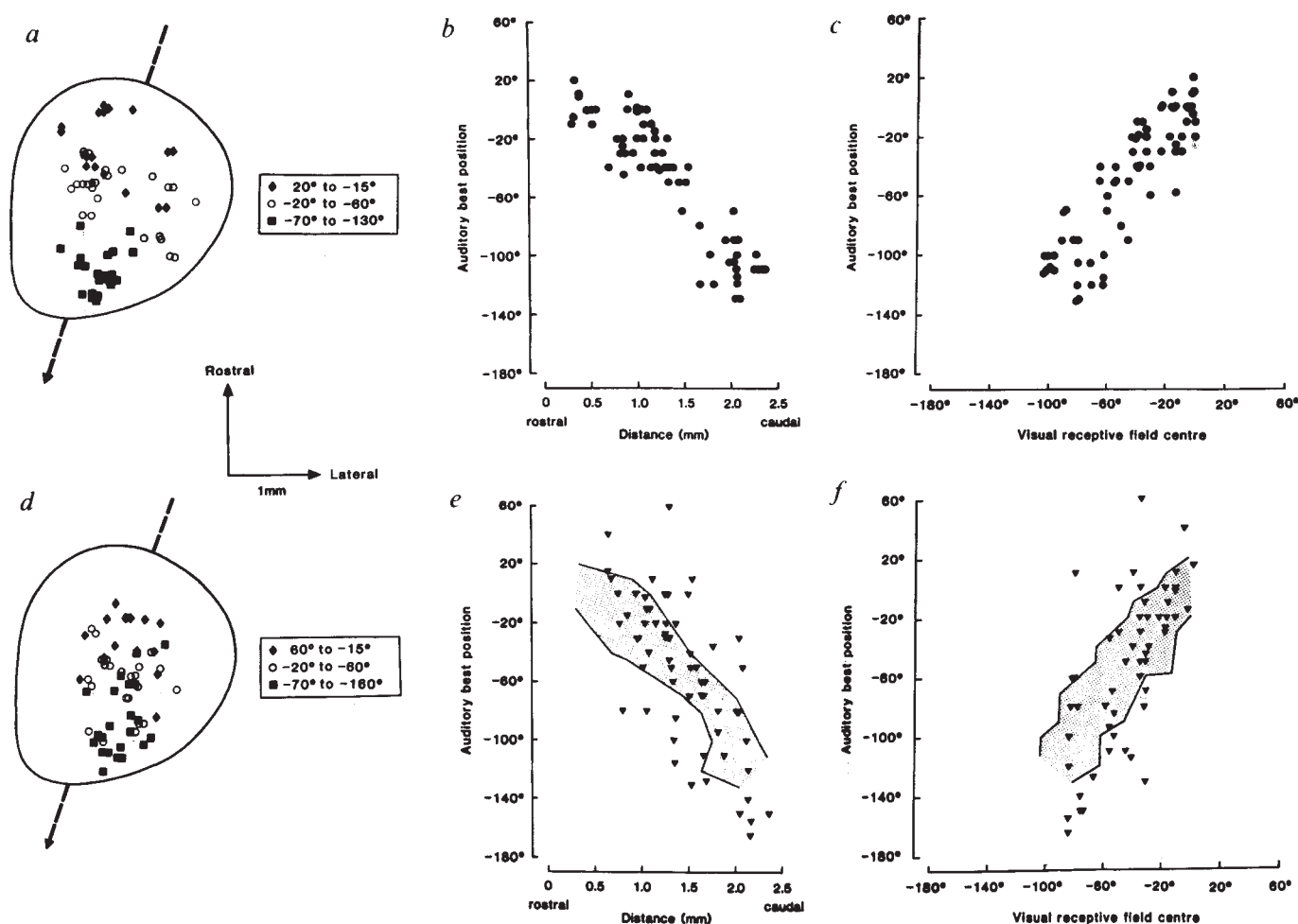


Fig. 1 *a*, Surface view of the right SC showing the distribution of auditory neurons recorded in the deep layers, for which a best position in azimuth could be ascribed. These data are pooled from 13 normal adult ferrets. Electrolytic lesions were placed in all electrode tracks, which were reconstructed from frozen, 50 μ m frontal sections of the SC cut at the same angle in every animal. The location of each unit was determined by calculating its depth below the surface, from the microdrive readings, and its distance along the rostrocaudal and mediolateral extents of the nucleus, which, because of the clear boundaries of the superficial region of the SC, could be determined to within one section. The coordinates of each unit, expressed as a proportion of the distance along the rostrocaudal and mediolateral axes, were then transposed to a standard surface representation of the SC. The auditory receptive fields were mapped with white-noise (0.5–40 kHz) bursts presented from a loudspeaker, which was moved using a hoop system²³ under free-field, anechoic conditions. The best position of a receptive field was defined as the speaker position for which the cell gave its largest response after subtraction of the average spontaneous spike count⁶. In most cases, this remained constant with changes in sound intensity. The different symbols represent the locations of cells that responded best to sounds within the ranges of positions indicated in the box. The dashed line runs orthogonal to the mean angle of isoazimuth contours derived by fitting regression lines through the locations of cells with the same best azimuthal position. The distance from the rostrolateral margin of the SC to each cell was measured in a direction parallel to the dashed line and plotted in *b* against best azimuthal position. *c*, The relationship between the auditory best positions of these cells and the centre of the visual receptive fields of cells in the overlying superficial layers of the SC that were recorded in the same electrode penetration. These were mapped using a discrete flashing light source attached to the hoop. *d*, Surface view of the right SC showing, as in *a*, the distribution of auditory units pooled from 11 ferrets that were reared with the ipsilateral ear occluded with an ear-mould impression compound from postnatal days 23–24. These ear plugs were renewed under anaesthesia (18 mg kg⁻¹ i.m. alphaxalone/alphadolone acetate) at approximately weekly intervals as the animals grew up. The attenuation caused by the plugs varied in a frequency-dependent way between 5 and 60 dB. The best positions of these units, measured with the ear plug still in place at sound intensities of more than 20 dB above threshold (intense enough to overcome the shadowing effects of the head so as to reach both external ears at all positions tested), are plotted against rostrocaudal distance in *e* and against visual receptive field centre in *f*. In both *e* and *f*, the stippled regions indicate the envelope of the normal data shown in *b* and *c*, respectively.

directions in space⁶. Preferred sound directions vary topographically in azimuth from 20° into the ipsilateral hemifield to 130° into the contralateral hemifield for cells located between the rostrolateral and caudomedial ends of the SC (Fig. 1*a, b*). They are closely aligned with the centres of visual receptive fields of cells recorded in the superficial layers during the same electrode penetration (Fig. 1*c*). Acute occlusion of the ear ipsilateral to the recording site during such an experiment results in a loss of spatial selectivity for most auditory units in the SC of normal adult guinea pigs¹⁴ and cats¹⁵, except for near-threshold stimuli. This suggests that the selectivity for suprathreshold sounds is dependent on binaural cues.

By contrast, in the SC of ferrets that had been reared with one ear chronically occluded, the receptive fields of auditory neurons were spatially tuned, with clearly-defined best positions at all sound intensities tested. A definite progression in the best

positions was apparent for a rostro-caudal sequence of electrode penetrations, although there was more variation than in normal animals (Fig. 1*d, e*) and a looser relationship between auditory and visual topography (Fig. 1*f*). For most units, removal of the ear plug during recording resulted in a shift in the auditory best position away from the side of the occluded ear by an average of about 30°. Thus, as previously demonstrated in the barn owl optic tectum^{10,11}, auditory spatial tuning and the topographic representation of auditory space in the ferret SC can adjust to distorted binaural cues during development.

To see whether the visual input to the SC provides a template for the maturation of the auditory map, we examined the effect of inducing an outward deviation of the eye by 15–20° (produced by section of the medial rectus muscle), which caused a corresponding overall shift in the map of visual space. Thus, each location in the SC now represented a position further into the

Fig. 2 a, Best position in azimuth plotted against rostro-caudal distance in the right SC of auditory neurons recorded from five adult ferrets in which a simple outward deviation (exotropia) of the contralateral eye had been surgically induced by removal of the medial rectus muscle, under halothane and nitrous oxide anaesthesia, on postnatal days 27–28. The lids of this eye were then re-closed by suturing with (resorbable) chromic collagen because they normally remain closed until ~30–32 days of age. The right eye was enucleated at the same time. The left eye opened spontaneously after 3–6 days. The degree of lateral deviation of the eye was estimated in conscious, hand-held ferrets at various times during development by plotting the position of the optic axis, which was determined by aligning the Purkinje–Sanson images with the centre of the pupil. Compared with normal ferrets of the same age, the average position of the axis was shifted by 15–20° in all five animals. The data shown here were obtained at sound intensities of more than 20 dB above threshold and were derived as described in Fig. 1. The stippled area shows the envelope of the normal data from Fig. 1b, and the thicker outline below is the same envelope displaced downwards (contralaterally) by the mean outward deviation of the eye (18°), as indicated by the arrows. A comparison of linear regression lines fitted through the data from the normal and strabismic ferrets revealed no difference between their slopes ($r = 0.71$; d.f. = 109; $P = 0.24$), but a highly significant difference between the positions of the lines (analysis of covariance: $t = 4.81$; d.f. = 110; $P < 0.0005$). **b**, Relationship between auditory best positions of deep-layer neurons and visual receptive-field centres of superficial layer neurons recorded in the same electrode penetrations for the five animals. We have not distinguished between different animals here because of the similarity in the shift in eye position in all five. The stippled region indicates the envelope of the normal data shown in Fig. 1c.

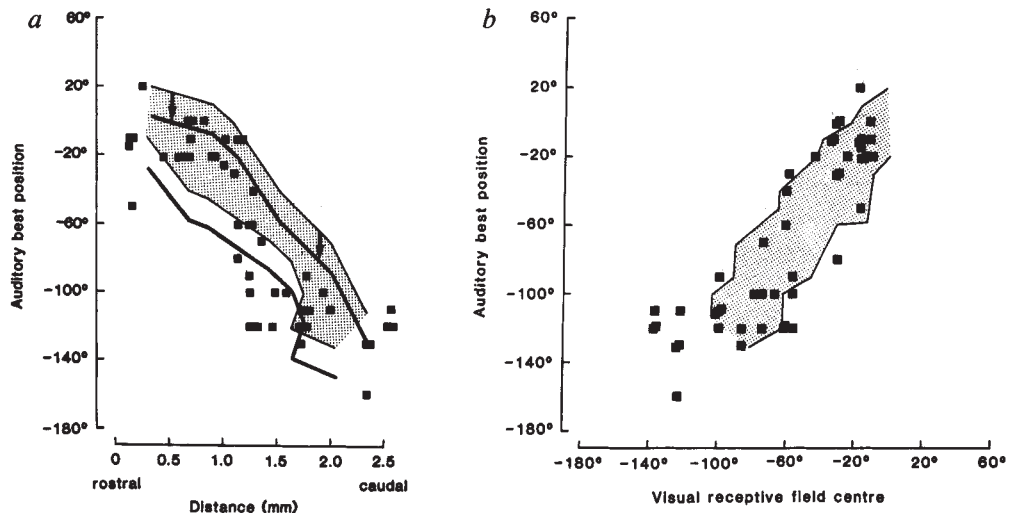
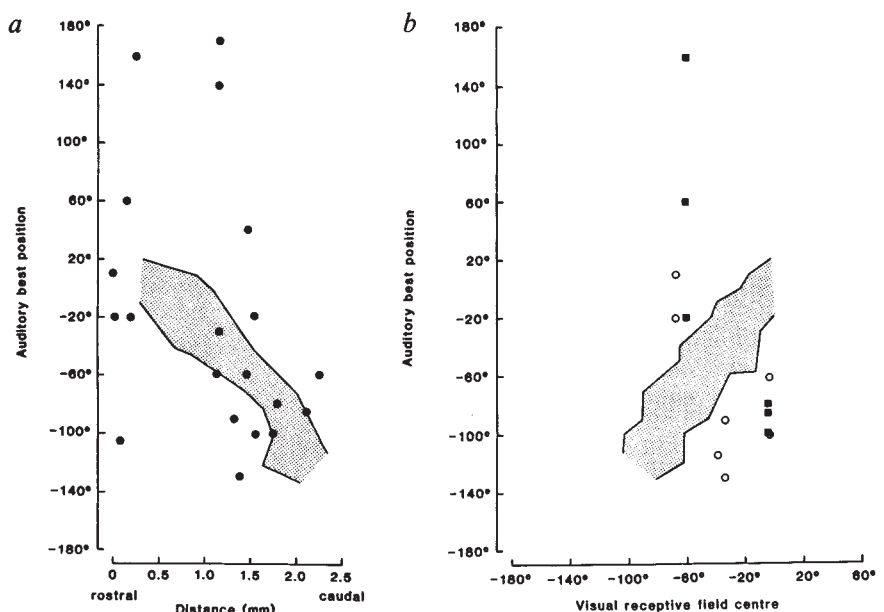


Fig. 3 a, best position in azimuth plotted against rostro-caudal distance in the right SC for auditory neurons recorded from five adult ferrets in which the contralateral eye had been rotated between postnatal days 25 and 27. Under halothane and nitrous oxide anaesthesia, the extraocular muscles of the left eye were sectioned and the globe gently rotated anticlockwise (incyclotorsion) by 180° around its anteroposterior axis. The eye-lids were temporarily sutured with chromic collagen and the right eye was enucleated. Some derotation of the left globe took place during this time, but, in all animals, the eye remained intorted by between 110° and 180°. No further derotation occurred and limited ocular mobility was observed, suggesting at least partial reattachment of the extraocular muscles. These data were obtained at sound intensities of more than 20 dB above threshold and were derived as described in Fig. 1. **b**, Relationship between auditory best positions of deep-layer neurons and visual receptive-field centres of neurons in the superficial layers of the same electrode penetration for two animals with different degrees of eye rotation (and therefore different orientations of the visual map). Open circles represent data from an animal with a 150° intorsion of the left eye and filled squares an animal with a 180° intorsion. The stippled regions in **a** and **b** indicate the envelope of the normal data shown in Figs. 1b and c, respectively.



contralateral visual field than normal. The auditory properties of cells in these animals appeared unchanged, except that the array of best positions was also displaced further contralateral compared with the normal range (Fig. 2a), by an amount approximately the same as the change in eye position. The entire auditory map was therefore shifted rostrally in the SC by about 300 μ m, so as to remain closely aligned with the visual map (Fig. 2b).

Neonatal eye rotation (by 110° to 180°) produced a corresponding rotation of the visual map in the contralateral SC (that is, the relationship between the retina and SC was unaffected).

Compared to all other groups of animals in this study, fewer acoustically-responsive cells were found in the SC of eye-rotated ferrets; whereas auditory cells were recorded in 85% of electrode tracks through the SC of normal ferrets, they were found in only 31% of tracks in the eye-rotated animals. Also many of these cells had abnormal firing patterns, and were broadly tuned for sound direction, responding to most loudspeaker positions tested. Where best positions could be ascribed, there was no indication of any topography in their distribution within the SC (Fig. 3a) and no correspondence with the rotated array of visual receptive fields (Fig. 3b).

Our results indicate that early acoustic experience plays an important role in the maturation of auditory receptive fields, as the mechanisms by which both the spatial tuning and the topographic map are generated can compensate for the effects of ear plugging in neonatal ferrets, but not in adult animals^{14,15}. The visual input to the SC, however, appears to be the major guiding influence in preserving the registration of the two maps during development: small displacements of visual topography (created by strabismus) cause a corresponding shift in the auditory map, but gross disturbance of the representation of the visual field (after eye rotation) exceeds the capacity of the mechanism for maintaining alignment and results in complete disorganization of the auditory map.

Even in the SC of adult primates¹⁶ (but not cats¹⁷), auditory receptive fields within the oculomotor range can shift during voluntary changes in eye position. Signals related to eye position might also play a part in the development of aligned sensory maps: the shift in the auditory map following strabismus and its disruption after eye rotation could conceivably be due to disturbances in signals about eye position reaching the SC¹⁸. We think it more likely, however, that the modified auditory responses that we have seen throughout the SC of the paralysed, anaesthetized ferret result primarily from the altered visual topography during a specific sensitive period. Eye rotation at a slightly later stage, just a few days after normal eye opening, does not change the topography of the auditory map, but merely increases the range of best positions represented at each location in the SC¹⁹. Thus, a limited amount of visual experience may be sufficient to stabilize the neural circuitry underlying the auditory representation.

Much evidence suggests that plasticity of the input from the two eyes on binocularly-driven, central neurons enables the system to adjust to changes in interocular relationships during growth²⁰⁻²². Similarly, as the head and pinnae increase in size, the correlation between the positions of sounds in space and the acoustic cues from which auditory receptive fields are derived will change. The cross-modal reorganization in receptive fields that we have demonstrated is therefore presumably a reflection of a normal plastic relationship between the visual and auditory systems, which compensates for growth-dependent changes, with the result that the alignment of sensory maps in the SC is maintained, thereby allowing both of them appropriate access to the motor maps found there.

We thank Michael Keating for comments on the manuscript, Nancy Kowalchuk for technical assistance and I. G. Vlachonikolis for statistical advice. Financial support was provided by the Lister Institute of Preventive Medicine, the MRC, SERC and the E. P. Abraham Research Fund.

Received 4 November 1987; accepted 14 January 1988.

1. Movshon, J. A. & Van Sluyters, R. C. *A. Rev. Psychol.* **32**, 477-522 (1981).
2. Kaas, J. H., Merzenich, M. M. & Killackey, H. P. *A. Rev. Neurosci.* **6**, 325-356 (1983).
3. *Development of Auditory and Vestibular Systems* (ed. Romand, R.) (Academic, New York, 1983).
4. King, A. J. & Palmer, A. R. *J. Physiol., Lond.* **342**, 361-381 (1983).
5. Middlebrooks, J. C. & Knudsen, E. I. *J. Neurosci.* **4**, 2621-2634 (1984).
6. King, A. J. & Hutchings, M. E. *J. Neurophysiol.* **57**, 596-624 (1987).
7. Knudsen, E. I., Esterly, S. D. & Knudsen, P. F. *J. Neurosci.* **4**, 1001-1011 (1984).
8. Knudsen, E. I., Knudsen, P. F. & Esterly, S. D. *J. Neurosci.* **4**, 1012-1020 (1984).
9. Knudsen, E. I. & Knudsen, P. F. *Science* **230**, 545-548 (1985).
10. Knudsen, E. I. *Science* **222**, 939-942 (1983).
11. Knudsen, E. I. *J. Neurosci.* **5**, 3094-3109 (1985).
12. Linden, D. C., Guillery, R. W. & Cucchiari, J. *J. comp. Neurol.* **203**, 189-211 (1981).
13. Moore, D. R. *Brain Res.* **253**, 309-311 (1982).
14. Palmer, A. R. & King, A. J. *Hearing Res.* **17**, 267-280 (1985).
15. Middlebrooks, J. C. *J. Neurophysiol.* **57**, 688-701 (1987).
16. Jay, M. F. & Sparks, D. L. *J. Neurophysiol.* **57**, 35-55 (1987).
17. Harris, L. R., Blakemore, C. & Donaghy, M. *Nature* **288**, 56-59 (1980).
18. Sparks, D. L. *Physiol. Rev.* **66**, 118-171 (1986).
19. King, A. J., Hutchings, M. E., Moore, D. R. & Blakemore, C. *Soc. Neurosci. Abstr.* **11**, 450 (1985).
20. Keating, M. J. *Br. med. Bull.* **30**, 145-151 (1974).
21. Pettigrew, J. D. in *Neuronal Plasticity* (ed. Cotman, C. W.) 311-330 (Raven, New York, 1978).
22. Blakemore, C. *Proc. R. Soc. B* **204**, 477-484 (1979).
23. Annett, C. S. J., Howse, B. P. A., Hutchings, M. E. & King, A. J. *J. Physiol., Lond.* **382**, 21P (1987).

A new type of imaging optics in compound eyes

D.-E. Nilsson

Department of Zoology, University of Lund, Helgonavägen 3, S-22362 Lund, Sweden

Many arthropods have compound eyes, made up of numerous separate visual units or ommatidia. In apposition eyes, the ommatidia are optically isolated from each other and provide a poor photon catch because the lenses are so minute¹. Much more efficient use of the eye's surface is obtained in superposition eyes, where many ommatidia cooperate to form a superimposed image on the retina¹. I report here that, contrary to previous belief², the eyes of many crabs and hermit-crabs work as superposition eyes by employing imaging optics of a conceptually new kind. Imaging is accomplished by a remarkable combination of ordinary lenses, cylindrical lenses, parabolic mirrors and light-guides. Despite the impressive complexity of the new mechanism, it is easy to see how this parabolic superposition eye must have evolved from an ordinary apposition compound eye.

Since the overall image in a compound eye is always upright, the individual ommatidium in a superposition eye must itself produce an upright image. Only two optical systems with this property have been described in compound eyes (Fig. 1a, b). The first was discovered nearly 100 years ago by Sigmund Exner³. He described a graded refractive index structure which acts as an astronomical telescope with two lenses, one reinverting the image produced by the other. This type is common in both insects and crustaceans^{2,4}. A second type of optics in superposition eyes was discovered by Vogt⁵ in 1975 and independently by Land⁶ a year later. They found that crayfish and shrimps use orthogonal sets of mirrors to produce the upright image on the retina. Both principles later found applications in optical engineering⁷; the graded index system is now used in some optical fibres, and the arrangement of orthogonal mirrors proved to be a useful design for X-ray telescopes.

In superposition compound eyes, the two designs are distinguished by the crystalline cones which are either circular in cross-section and contain a powerful refractive index gradient (refracting superposition; Fig. 1a) or square in cross-section and operate by internal reflections (reflecting superposition, Fig. 1b)^{1,2,8}. I have now found that many crabs (*Brachyura*) and hermit-crabs (*Anomura*), which were previously believed to have simple apposition eyes, in fact have well focused superposition eyes (Fig. 2a, b). However, the design does not conform with either of the two known superposition principles. Instead it turns out that these animals have a third, previously unknown type of superposition optics, here termed parabolic superposition (Fig. 1c).

I first encountered the new principle in portunid crabs of the genus *Macropipus*. In the dark-adapted state these eyes have a wide pigment-free zone separating the crystalline cones from the retina. This clear zone is traversed by light-guides formed by narrow columns of the retinula cells. The crystalline cones do not contain refractive index gradients nor are they square in cross-section. To explain how the erect image is formed, we have to consider the ray path first in a section along the ommatidia (Fig. 3a, d), and then as viewed from a point on the optical axis (Fig. 3b, c).

The corneal lenses are powerful and the focal distance corresponds approximately to the distance down to the pointed end of the crystalline cone. An oblique bundle of rays will be reflected inside the cone at its border. These reflections are augmented by a multilayer coating of regularly arranged endoplasmic reticulum (Fig. 2c). Seen from the side, the middle section of the cone has an inward parabolic curvature which cancels the

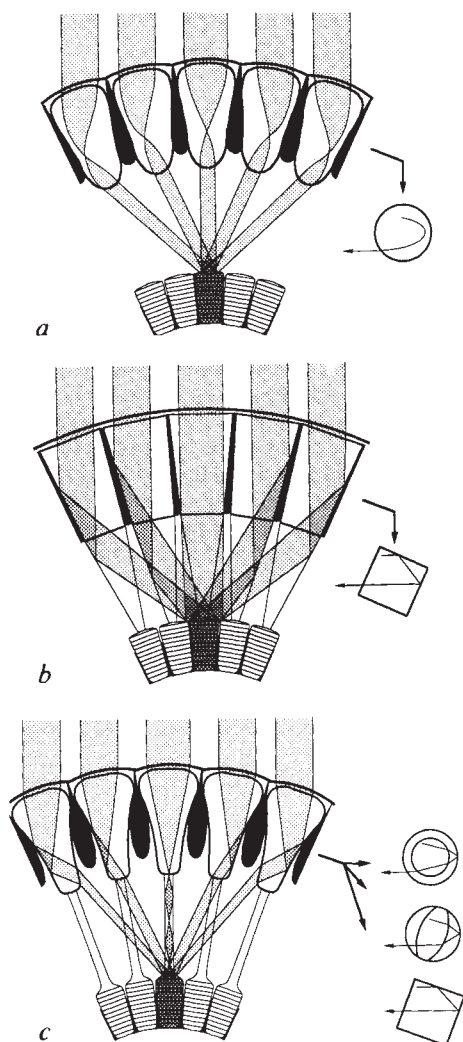


Fig. 1 Schematic diagrams of compound eyes forming a superposition focus of rays from a distant point-source. All superposition eyes share the same basic feature of an unpigmented zone separating crystalline cones from rhabdoms. *a*, The refracting superposition type employs a refractive index gradient in the crystalline cones, with highest optical density along the axis. Rays are bent smoothly and the top view of a crystalline cone reveals an elliptical ray-path. The optical system in each crystalline cone is that of an astronomical telescope. *b*, The reflecting superposition type instead relies on reflections and no lenses are necessary. In a top view the crystalline cones are square and operate on the same principle as a corner-reflector. *c*, The new superposition principle, here termed parabolic superposition because, in a side view, the reflecting borders of the crystalline cones need be parabolic to cancel the convergence of rays caused by the corneal lenses. The new principle is further characterized by light-guides crossing the clear-zone and by the crystalline cones containing an internal cylindrical lens (upper two cross-sections) or being square in cross-section.

convergence of rays so that the reflected beam is recollimated and directed to a rhabdom coaxial with the light source (Figs. 1c, 3a). This lens/mirror combination is equivalent to a positive lens followed by a negative one, such as in a Galilean telescope.

The most interesting feature of the optical system is seen in a top view of the crystalline cone. The reflecting border of the cone is circular in cross-section and on its own it would ruin the collimation of the beam. To prevent this, the cone contains an internal cylindrical lens (Fig. 2d) which shortens the focal length so that the focus is reached exactly when the beam hits the reflecting border (Fig. 3b). From this point, the diverging fan of reflected rays is recollimated as it again has to cross the cylindrical lens to emerge into the clear zone (Fig. 3c). Note

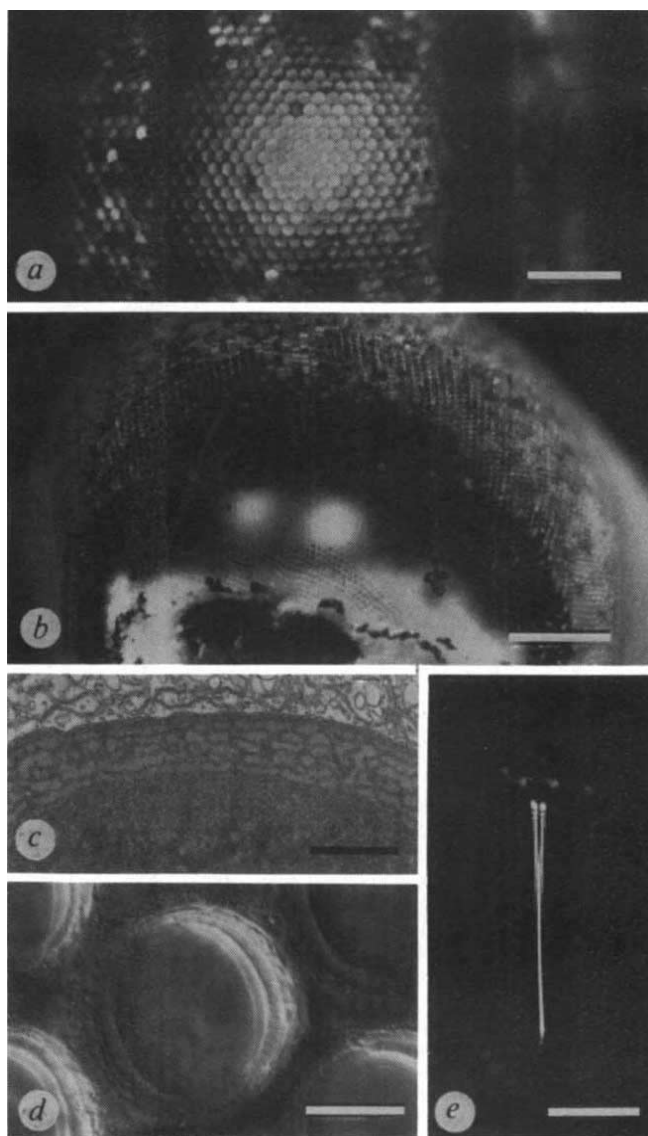


Fig. 2 Micrographs demonstrating some features of the parabolic superposition eye of a swimming crab (*Macropipus depurator*). *a*, Due to a reflecting tapetum below the retina, a dark-adapted eye displays an eye-shine upon illumination from the direction of observation. Here, the cornea is illuminated by vertical stripes but the eye-shine is seen equally well in non-illuminated areas between stripes. This demonstrates that the ommatidia are not optically isolated and thus the eye cannot be of apposition type. Scale bar: 0.25 mm. *b*, An eye is illuminated by two point-sources, 10° apart, and a hole is cut at the side of the eye so that the image on the retina can be inspected directly. The image quality is comparatively good and the two point-sources are clearly resolved. Scale bar, 0.5 mm. *c*, Electron micrograph of the border of a cross-sectioned crystalline cone, showing the multilayer reflector formed by regularly arranged endoplasmic reticulum. The layer thickness of ~ 90 nm is that expected from a $1/4$ wavelength interference reflector¹³ for green light in a medium with a refractive index of 1.4. The reflector is circular in cross-section but reveals a near parabolic curvature in longitudinal sections. Scale bar, 1 μ m. *d*, Cross-section of crystalline cones cut from a frozen eye and observed with Nomarski optics. The internal cylindrical lens stands out because of its higher refractive index. Scale bar, 15 μ m. *e*, An isolated crystalline cone illuminated obliquely with a solid cone of light from a HeNe laser. The illumination arrangement mimics the input from an oblique point-source focused by the corneal lens. The internal cylindrical lens shortens the focal length so that an image falls on the cone border. Since the lens is cylindrical, the focal length is only shortened in one plane and the point-source is imaged as a line along the cone border. Most of the light is reflected but some is scattered, making the image visible. Scale bar, 30 μ m.

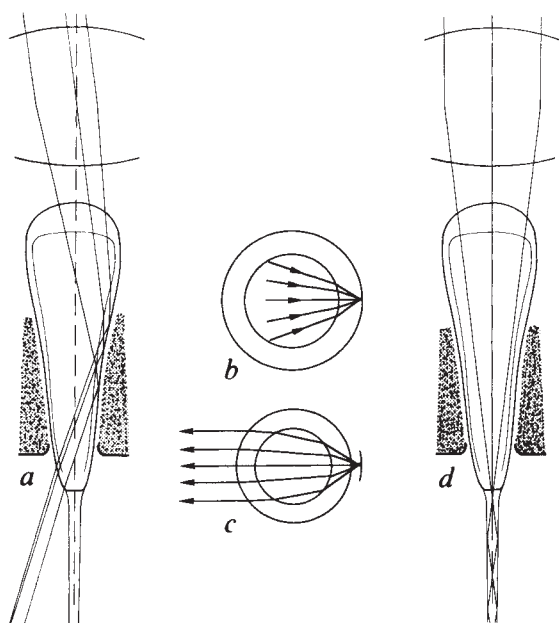


Fig. 3 Ray paths in parabolic superposition eyes. *a*, Ray tracing through an ommatidial model of the crab *Macropipus depurator*. The model is based on Nomarski micrographs of fixed¹⁰ and isolated but otherwise untreated cones. Refractive indices were measured with an interference microscope. Parallel rays are incident at an angle of 8° to the axis. The convergence of rays caused by the corneal lens is cancelled by the parabolic curvature of the reflecting border so that a recollimated pencil of rays emerges from the cone. *b*, *c*, Schematic diagrams of the ray path seen in a top view of the cone. The simple geometry is of the type found in the hermit-crab *Eupagurus bernhardus*. Seen in this plane, the internal cylindrical lens shifts the focus to a point on the reflecting border which is circular in cross-section (*b*). After the reflection and farther down in the cone (*c*), the diverging bundle of rays becomes recollimated by the cylindrical lens. *d*, Paraxial incident light, which does not reach the reflecting border, is focused at the proximal tip of the cone and then trapped in a light-guide which leads it through the clear zone (here shown in the *Macropipus* model). The 4 µm wide light-guide has a refractive index of 1.37 and is surrounded by haemolymph with a refractive index of 1.34. This makes an acceptance cone of 24° which is just right to include all direct rays from the cornea but to let the oblique reflected rays escape.

that the internal lens is cylindrical and therefore does not add any focusing power to the ray path as seen from the side (see Fig. 2*e*). For each individual ray in the side-view (Fig. 3*a*) there is thus a whole set of rays in the top view (Fig. 3*b*, *c*). In effect, there are two distinct types of optical systems coexisting and operating in different planes.

Paraxial incident rays are handled in yet another way. These rays will not hit the reflecting border and would diverge strongly if they were to emerge freely from their focus at the cone tip. It is here that the light-guides come in. Their function is to trap axial light and guide it through the clear zone down to the rhabdom (Fig. 3*d*).

There are many variations on the basic theme of parabolic superposition. The common hermit-crab *Eupagurus bernhardus* has, in part of its eye, exactly the design described above. Portunid and majid crabs differ slightly by having a bipartite cylindrical lens (Figs 1*c*, 2*d*)—probably to minimize aberrations at the periphery of the cylindrical surface. The xanthid crabs (for example, *Trapezia*), have no cylindrical lens at all inside the crystalline cone, but use a square cross-section in the same way as in reflecting superposition eyes (see Fig. 1*b*).

Despite the large variation, four criteria are common to all parabolic superposition eyes: (1) the crystalline cones have

inward parabolic curvature along part of the length; (2) corneal lenses are focused at the distal end of the clear zone; (3) the clear zone is crossed by light-guides; (4) the crystalline cones are either circular in cross-section with an internal cylindrical lens, or square in cross-section without internal structures. The above characteristics hold for a large number of true crabs, hermit-crabs and, interestingly, also for a few may-flies⁹.

The light-guides that are required to take care of axial light are often a narrower distal part of the rhabdom and in some species such as *Carcinus* and *Hyas*, the superposition mechanism involves the cooperation of so few ommatidia that the eyes approach a pure apposition design. Contrary to the case with refracting^{10,11} and reflecting¹² superposition eyes, there are no obvious non-functional intermediates involved in the evolutionary transformation from apposition to parabolic superposition. Even the most well-developed parabolic superposition eyes employ an apposition mechanism for axial rays, and the fact that almost every conceivable state of development of a parabolic superposition mechanism is present today indicates that it can evolve easily from a simple apposition eye.

The discovery of this new but common type of eye now calls for a survey of the eyes of crabs, hermit-crabs and may-flies so that the mechanism can be worked out in more detail and its efficiency compared with conventional types of optics.

Received 26 October 1987; accepted 20 January 1988.

- Land, M. F. in *Handbook of Sensory Physiology* Vol. VII/6B (ed. Autrum, H.) 471–592 (Springer, Berlin, 1981).
- Fincham, A. A. *Nature* **287**, 729–731 (1980).
- Exner, S. *Die Physiologie der facettierten Augen von Krebsen und Insekten* (Deuticke, Leipzig, 1891).
- Land, M. F. in *Comprehensive Insect Physiology Biochemistry and Pharmacology* (eds Kerkut, G. A. & Gilbert, L. I.) 225–275 (Pergamon, Oxford, 1985).
- Vogt, K. Z. *Naturforschung* **30c**, 691 (1975).
- Land, M. F. *Nature* **263**, 764–765 (1976).
- Land, M. F. *Nature* **287**, 681–686 (1980).
- Land, M. F. in *Photoreception and Vision in Invertebrates* (ed. Ali, M. A.) 401–437 (Plenum, New York, 1984).
- Horridge, G. A., Marčelja, L. & Jahnke, R. *Proc. R. Soc. B* **216**, 25–51 (1982).
- Nilsson, D.-E., Land, M. F. & Howard, J. *Nature* **312**, 561–563 (1984).
- Nilsson, D.-E., Land, M. F. & Howard, J. *J. comp. Physiol.* (in the press).
- Nilsson, D.-E. *Nature* **302**, 818–821 (1983).
- Miller, W. H. in *Handbook of Sensory Physiology* Vol. VII/6A (ed. Autrum, H.) 69–143 (Springer, Berlin, 1981).

A new natriuretic peptide in porcine brain

Tetsuji Sudoh, Kenji Kangawa, Naoto Minamino & Hisayuki Matsuo*

Department of Biochemistry, Miyazaki Medical College, Kiyotake, Miyazaki 889-16, Japan.

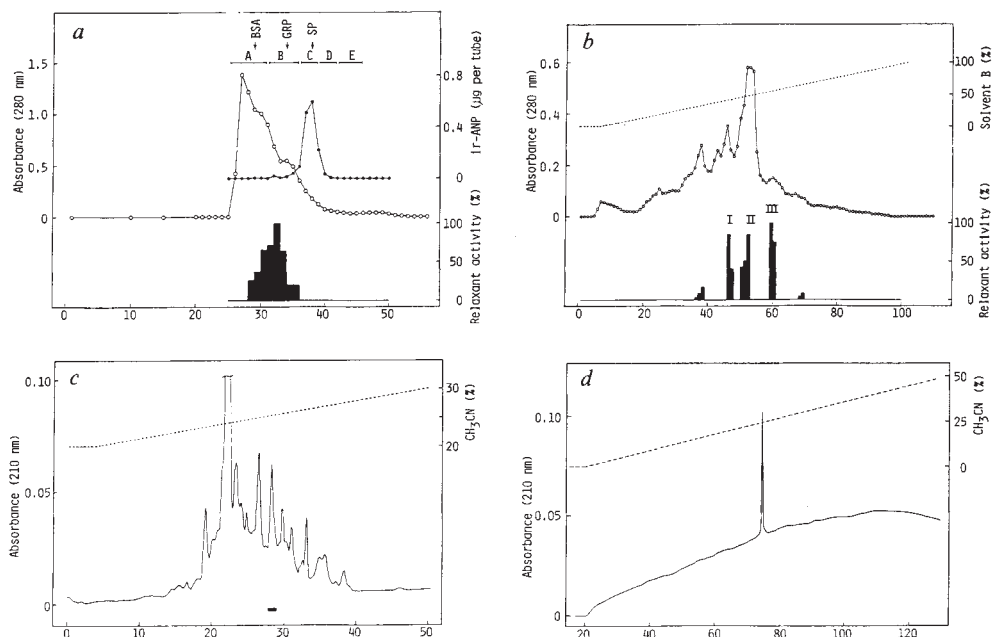
Atrial natriuretic peptide (ANP), a hormone secreted from mammalian atria, regulates the homeostatic balance of body fluid and blood pressure^{1–3}. ANP-like immunoreactivity is also present in the brain, suggesting that the peptide functions as a neuropeptide^{4,5}. We report here identification in porcine brain of a novel peptide of 26 amino-acid residues, eliciting a pharmacological spectrum very similar to that of ANP, such as natriuretic-diuretic, hypotensive and chick rectum relaxant activities. The complete amino-acid sequence determined for the peptide is remarkably similar to but definitely distinct from the known sequence of ANP^{1,3,6}, indicating that the genes for the two are distinct. Thus, we have designated the peptide 'brain natriuretic peptide' (BNP). The occurrence of BNP with ANP in mammalian brain suggests the possibility that the physiological functions so far thought to be mediated by ANP may be regulated through a dual mechanism involving both ANP and BNP.

* To whom correspondence should be addressed.

Fig. 1 Purification of pBNP from porcine brain extracts. *a*, Gel filtration on Sephadex G-25(fine) of a portion of M_r 2–5K obtained by a gel filtration on Sephadex G-50 of SP-III fraction. Column size, 4.5×140 cm; flow rate, 40 ml h^{-1} ; eluent, 1 M AcOH (acetic acid); fraction size, 40 ml per tube. Size markers used: bovine serum albumin (BSA); gastrin-releasing peptide (GRP); substance P (SP). Relaxant activity on chick rectum is indicated by black bars. ANP-like immunoreactivity emerged in fraction C, *b*, Cation exchange chromatography of fraction B (tubes 32–36 in *a* yielded 410 mg). Column, CM-52 (2.4×45 cm; Whatman); flow rate, 33.5 ml h^{-1} ; fraction size, 20 ml per tube; solvent system, linear gradient elution from A to B. A, 10 mM HCO_2NH_4 (pH 6.6): CH_3CN , 90:10; B, 0.5 M HCO_2NH_4 (pH 6.6): CH_3CN , 90:10 (v/v). Rectum relaxant activity was separated into three major peaks: I (tubes 46–47), II (tubes 51–53) and III (tubes 60–61). Peptides (71 mg) in peak II were re-purified by cation exchange HPLC (TSK CM-2SW, Tosoh, 7.6×300 mm) to give a single bioactive peak (tubes 39–41). The active peak obtained⁷ was subjected to RP-HPLC on a Chemcosorb 7-diphenyl column (4.6×250 mm; Chemco) to yield a single rectum-relaxant peak emerging at 43–46 min. Solvent system, linear gradient elution from A to B for 120 min. H_2O : CH_3CN :10% TFA (trifluoroacetic acid) for A was 90:10:1, for B was 40:60:1 (v/v). *c*, RP-HPLC of the rectum-relaxant peak obtained by RP-HPLC on a 7-diphenyl column. Activity emerged in a well separated peak at 27.8–29.0 min. Column, Chemcosorb 3-ODSH (8×75 mm; Chemco), flow rate, 2.0 ml min^{-1} ; solvent system, linear gradient elution from A:B=80:20 to A:B=0:100 for 192 min. H_2O : CH_3CN :10% TFA for A was 90:10:1, for B was 40:60:1 (v/v). Final purification of pBNP was performed by RP-HPLC on a semi-micro column of Hitachi 3063 (1.2×150 mm) to yield pBNP eluting at 45.2–46.4 min in a pure state. The yield was about 1.5 μg starting from 200 pig brains. Chromatographic conditions were identical to those described in *d*, except that a shallow linear gradient from A:B=70:30 to A:B=45:55 for 60 min was used. *d*, RP-HPLC on a semi-micro column of a portion (~ 50 ng) of purified pBNP gave a single peptide peak eluting at 74.5 min, confirming the homogeneity of the peptide. A synthetic pBNP also comigrates with natural pBNP under identical conditions. Column: Hitachi 3063 (1.2×150 mm); flow rate: $50 \mu\text{l min}^{-1}$; solvent system, linear gradient elution from A to B for 135 min. H_2O : CH_3CN :10% TFA for A was 100:0:1, for B was 40:60:1 (v/v).

Methods. Purification was carried out with acid extracts of porcine brain, using a similar method to that described in ref. 14. Chick rectum relaxant activity in the extracts was monitored. Porcine brain (18.6 kg obtained from 200 pigs) was minced and heat-treated at 95°C for 10 min in two volumes (v/v) of water to inactivate intrinsic proteases. After cooling, a final concentration of the aqueous layer was made to 1 M AcOH–20 mM HCl, by adding HCl and AcOH, and boiled tissues were extracted by homogenizing with a Polytron mixer. After condensation and acetone-precipitation (75%), the supernatant was evaporated to dryness. Dry materials were dissolved in 1 M AcOH and submitted to batchwise chromatography on SP-Sephadex C-25(H^+ -form) in the presence of 1 M AcOH. Successive elutions with 1 M AcOH, 2 M pyridine and 2 M pyridine–AcOH (pH 5.0) yielded three fractions of SP-I, SP-II and SP-III, respectively. SP-III fraction containing basic peptides (~ 6.0 g) was loaded on a preparative column of LC-SORB-ODS (Chemco) in the presence of 0.1 M AcOH. The materials adsorbed on the column were eluted with a solution of H_2O : CH_3CN :10% TFA at 40:60:1 to yield 5.1 g of basic peptides, which were separated by gel filtration on Sephadex G-50. A portion corresponding to M_r 2–5K was collected and subjected to further purification. Appropriate amounts of samples were dissolved in 100–200 μl of 0.9% saline and used for each bioassay. Chick rectum relaxant activity was assayed by the described method¹, using strips of freshly isolated chick rectum⁷. Natriuretic and diuretic activities were assayed as described previously^{1,15}, after injection into assay rats through the jugular vein in one shot. Urine excretion and concentration of Na^+ , K^+ and Cl^- in urine were determined as described previously^{1,11}. Systemic blood pressure was measured from the carotid artery in rats (body weight of 300–400 g) anaesthetized with pentobarbital^{1,11}. Indicated amounts of peptides were administered through a cannula in the femoral vein. Radioimmunoassay for ANP was performed using an antiserum specific for α -hANP^{12,13}. The antiserum shows 20% crossreactivity to rat α -ANP, which has a single replacement of Met-to-Ile, indicating that epitope recognized includes position 12. Crossreactivity of the antiserum with synthetic pBNP is less than 0.1%.

Porcine brain natriuretic peptide (pBNP) was isolated from acid extracts of porcine brain, by monitoring its potent relaxant effect on chick rectum⁷. This assay was also used for assessment of ANP in mammalian atrial tissues^{1,2}. A portion of the extracts having low relative molecular mass (M_r) was subjected to cation exchange chromatography to collect a fraction (SP-III) containing basic peptides, which was applied to reverse-phase high pressure liquid chromatography (RP-HPLC) on a preparative C-18 column followed by gel filtration on Sephadex G-50. A fraction of M_r 2–5K thus collected, was again separated on Sephadex G-25. Remarkably rectum-relaxant activity was observed in tubes 28–36 (Fig. 1a, fractions A and B). From ANP-like immunoreactivity emerging in fraction C, α -ANP-[4–28] and [5–28] were identified as the brain forms of porcine ANP⁸. They were not detected by the bioassay, however, as fraction C overlapped with rectum stimulant activity. Fraction B (tubes 32–36) was further separated by cation exchange



chromatography on a CM-52 column into three major peaks of activity (Fig. 1b). We purified pBNP from peak II (tubes 51–53); the rectum-relaxant activities in peaks I and III were identified as CGRP (calcitonin gene-related peptide) and VIP (vasoactive intestinal peptide), respectively. Subsequent cation exchange HPLC on CM-2SW of peak II yielded a single bioactive peak, from which pBNP was purified to homogeneity by successive RP-HPLC (Fig. 1c, d). The amino-acid composition of RCM (reduced and carboxymethylated)-pBNP was: CmCys, 1.62(2); Asp, 3.00(3); Ser, 2.56(3); Gly, 4.88(5); Val, 1.00(1); Ile, 0.95(1); Leu, 4.06(4); Tyr, 1.08(1); Phe, 0.98(1); Arg 5.02(5), indicating that pBNP is composed of 26 amino-acid residues, including two Cys residues. Based on these data, the isolation yield of pBNP was estimated to be about 1.5 μg (600 pmole), starting from 18.6 kg of brain tissue (200 pigs). Purified pBNP was sequenced by a gas-phase automated sequencer after conversion to RCM-pBNP to elucidate the complete amino-acid sequence

of pBNP (Fig. 2). Positive confirmation of the structure deduced above was performed by chromatographic comparison of natural pBNP and an identically sequenced synthetic peptide containing an intramolecular disulphide linkage. This unambiguously established the complete amino-acid sequence of pBNP. As identical peptide sequence is not found in the known peptide sequences stored in the data bank at PRF-SEQDB (Protein Research Foundation, Osaka, Japan), we conclude that pBNP is a new peptide. However, the peptide does show a remarkable homology to the known sequence of porcine α -ANP⁶, which is identical to human α -ANP(α -hANP)¹ (Fig. 2b). The highest homology was observed when pBNP was compared with α -ANP[4-28], which is one of the major brain forms of porcine ANP⁸. The 26 amino-acid residues of pBNP have seven amino-acid replacements and one insertion of (Arg) compared to α -ANP[4-28]. The 17 amino-acid sequence (α -ANP[7-23]) flanked by two Cys residues, thought to be essential for ANP activity, is highly conserved in the molecule of pBNP, although four residues in this region are replaced. This may explain the fact that pBNP does not crossreact with anti- α -hANP-antibody¹³, which is only 20% crossreactive even with rat α -ANP(α -rANP), which has only a single replacement (Met-to-Ile) at position 12 (see Fig. 1, legend). In addition, the C-terminal Arg-Tyr sequence of α -ANP, which is also important for ANP activity, is retained at the C-terminus of pBNP. As was suggested by the structural resemblances between ANP and BNP, pBNP was found to elicit a pharmacological spectrum very similar to α -hANP¹ (Table 1). Intravenous injection of

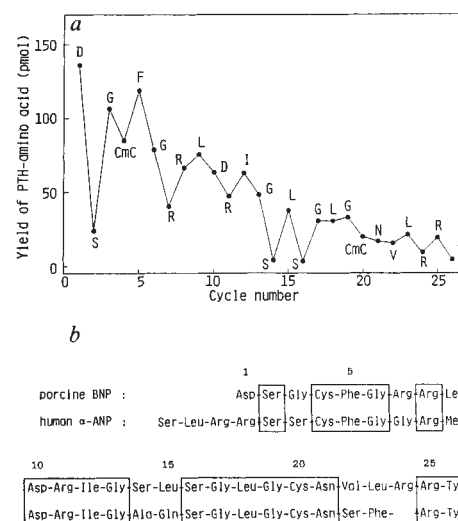


Fig. 2 Complete amino-acid sequence of porcine BNP compared with human α -ANP. *a*, Yield of phenylthiohydantoin (PTH)-amino acids liberated at each cycle of Edman degradation of RCM-pBNP; the one-letter amino-acid notation is used, except Cmc for carboxymethyl-cysteine. *b*, Comparison of porcine BNP with human α -ANP, which is common to higher mammals except rodent species. Intramolecular disulphide linkage is present between two Cys residues in each peptide. Identical residues are boxed.

Methods. After reductive S-carboxymethylation of cysteine residues performed by the described method^{1,11}, the resulting RCM-pBNP (~200 pmol) was sequenced with a gas-phase sequencer (Applied Biosystems Model 470A), coupled with HPLC identification of the resulting PTH-amino acids, as described^{1,11}. The amino-acid sequence thus determined for pBNP was verified by co-chromatography with synthetic BNP on RP-HPLC under the conditions as described in Fig. 1d, before and after reductive S-carboxymethylation of an intramolecular disulphide bond. Synthetic pBNP was made by solid-phase synthesis conducted on a chloromethylated polystyrene resin, as in the previous synthesis of α -hANP¹ with a slight modification. S-Acm-Boc-Cys was used for coupling of Cys residues. An intramolecular disulphide linkage of pBNP was formed by the action of iodine in the presence of AcOH¹⁶.

Table 1 Diuretic-natriuretic and hypotensive responses induced by pBNP and α -hANP

Dose (nmol)	0.1	pBNP 0.2	0.4	α -hANP 0.4
% Increase in urine output	121 $\pm$ 27	227 $\pm$ 42	368 $\pm$ 110	315 $\pm$ 52
Na ⁺ excretion	133 $\pm$ 4	376 $\pm$ 311	550 $\pm$ 291	422 $\pm$ 83
K ⁺ excretion	120 $\pm$ 16	210 $\pm$ 95	254 $\pm$ 85	278 $\pm$ 48
Cl ⁻ excretion	154 $\pm$ 44	274 $\pm$ 60	497 $\pm$ 161	451 $\pm$ 78
Decrease in mean blood pressure (mmHg)	5.4 $\pm$ 2.4	12.4 $\pm$ 3.2	19.2 $\pm$ 4.8	20.4 $\pm$ 5.0

Diuretic-natriuretic responses are expressed as % change (mean $\pm$ s.e.m.) in urine output, and in excretions of Na⁺, K⁺ and Cl⁻ from 15-min urine samples collected before and after i.v. injection into anaesthetized rats. (Seven rats were used for each sample.) Hypotensive responses are expressed as decrease (mean $\pm$ s.e.m.) in mean blood pressure. (Five rats were used for each sample.)

synthetic pBNP into anaesthetized assay rats resulted in a remarkable increase in excretion of urine and electrolytes into urine, in a manner very similar to that elicited by α -hANP. Furthermore, injection of pBNP into anaesthetized rats caused a significant decrease in mean blood pressure, comparable to that induced by α -hANP at the same doses. Thus, pBNP is as potent as α -hANP in hypotensive and diuretic-natriuretic activity. Note that the rectum-relaxant activity of pBNP is 3-4 times more potent than that of α -hANP, suggesting a unique feature in the physiological significance of pBNP. Our identification of BNP with ANP in porcine brain shows that a family of at least two natriuretic peptides exists in mammals, implying that physiological functions, such as water intake and salt appetite, so far thought to be regulated by ANP, may be controlled in a dual manner by ANP and BNP, at least in brain. Comparing isolation yields, the present study indicated that brain concentration of pBNP is $\sim 3 \times$ higher than that of ANP in pig brain.

As a number of neuropeptides originally identified in brain have been found often in the peripheral organs and *vice versa*, it is probable that BNP is also present in other organs, such as

heart, where it may function in concert with ANP to maintain the homeostatic balance of body fluid. ANP in higher mammals is initially synthesized as a 151-residue precursor^{3,9,10}, from which a 126-residue peptide is processed (γ -ANP)^{3,11} and stored in atrial granules¹². The γ -ANP is then converted to the 28-residue α -ANP, which circulates in the blood¹³. In brain, however, further processing of the precursor takes place to yield N-terminally shortened forms of α -ANP, such as α -ANP[4-28] and [5-28]^{5,8}. Taking such a tissue-specific diversity of processing into consideration, shorter or longer BNP-related peptides are probably present in other organs.

Finally, as the sequence of pBNP is not found in the known sequence of the ANP precursor³, the peptide is probably processed from its own precursor, genetically distinct from the ANP precursor. Furthermore, of the seven amino-acid changes observed between ANP and BNP, Arg to Asp at position 1 and Ser to Val at position 22 are not convertible by a single nucleotide substitution, indicating that the genes encoding the two molecules diverged at an early stage.

This work was supported in part by research grants from the Ministry of Education, Science and Culture of Japan.

Received 29 September 1987; accepted 8 January 1988.

- Kangawa, K. & Matsuo, H. *Biochem. biophys. Res. Commun.* **118**, 131-139 (1984).
- de Bold, A. J. *Science* **230**, 767-770 (1985).
- Matsuo, H. & Nakazato, H. *Endocrinology Metabolism Clinics of North America*, Vol. 16, *Atrial Natriuretic Factor* (eds Rosenblatt, M. & Jacobs, J. W.) 43-61 (Saunders, Philadelphia 1987).

4. Samson, W. K. *Endocrinology & Metabolism Clinics of North America*, Vol. 16, *Atrial Natriuretic Factor* (eds Rosenblatt, M. & Jacobs, J. W.) 145-161 (Saunders, Philadelphia, 1987).
5. Nakao, K. et al. *J. Hypertension* 4, (suppl. 6) 492-496 (1986).
6. Forssmann, W. G. et al. *Cell Tissue Res.* 238, 425-430 (1984).
7. Currie, M. G. et al. *Science* 221, 71-73 (1983).
8. Ueda, S. et al. *Biochem. biophys. Res. Commun.* 149, 1055-1062 (1987).
9. Oikawa, S. et al. *Nature* 309, 724-726 (1984).
10. Oikawa, S. et al. *Biochem. biophys. Res. Commun.* 132, 892-899 (1985).
11. Kangawa, K., Fukuda, A. & Matsuo, H. *Nature* 313, 397-400 (1985).
12. Miyata, A., Kangawa, K., Toshimori, T. & Matsuo, H. *Biochem. biophys. Res. Commun.* 129, 248-255 (1985).
13. Miyata, A., Toshimori, T., Kangawa, K. & Matsuo, H. *Biochem. biophys. Res. Commun.* 142, 461-467 (1987).
14. Minamino, N., Kangawa, K. & Matsuo, H. *Biochem. biophys. Res. Commun.* 130, 1078-1085 (1985).
15. de Bold, A. J., Borenstein, H. B., Veress, A. T. & Sonnenberg, H. *Life Sci.* 28, 89-94 (1981).
16. Kamber, B. et al. *Helv. chim. Acta* 63, 899-915 (1980).

Antigen chimaeras of poliovirus as potential new vaccines

Karen L. Burke, Glynis Dunn*, Morag Ferguson*,
Philip D. Minor* & Jeffrey W. Almond

Department of Microbiology, University of Reading, London Road, Reading RG1 5AQ, UK

* National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Potters Bar, Hertfordshire EN6 3QG, UK

Polioviruses occur as three distinct serotypes, 1, 2 and 3, and are composed of a single-stranded positive-sense RNA genome of ~7,450 nucleotides enclosed in an icosahedral particle of diameter 27 nm. The three-dimensional crystallographic structure of poliovirus type 1 has been determined at 2.9 Å resolution, providing a detailed knowledge of the folding and arrangement of the individual virus proteins, VP1-VP4¹. From this and the characterization of monoclonal antibody-resistant mutants²⁻⁷, the amino acids contributing to antigenic sites have been identified and located on the surface of the virus particle. Here we describe the construction and characterization of a poliovirus chimaera having a defined region of type 3 inserted into type 1. This virus has composite antigenicity and the substitute site is immunogenic in small animals and primates. The ability to construct such viruses has implications for the design of improved poliovirus vaccine strains or vaccines against other picornaviruses, such as hepatitis A.

The live attenuated poliovirus vaccines in current use are derived from strains developed by multiple passage of wild-type viruses in non-human primates and in tissue culture⁸. Although the use of these vaccines has almost eliminated poliomyelitis in developed countries, both type 2 and type 3 Sabin vaccine strains have been associated with a low level of poliomyelitis in vaccinees⁹. In contrast no well-documented cases of poliomyelitis have been attributed to the Sabin type 1 vaccine⁹. It has been shown that the type 1 vaccine has multiple attenuating mutations as compared with its neurovirulent progenitor¹⁰ whereas the type 3 strain has only two^{11,12}, providing an explanation for the relative stability of the Sabin type 1 vaccine strain.

Recent advances in poliovirus molecular biology and genetics

Table 1 Neutralization of chimaeric virus by monospecific polyclonal antisera

Antiserum	Sabin1	Sabin 2	Sabin 3	S1/3.10
Type 1	>5.75	0.0	0.0	>6.5
Type 2	0.0	>5.25	0.0	0.0
Type 3	0.0	0.0	>5.25	4.0

Reductions in titre ($\log_{10}$) of the infectious virus were determined by challenging 10-fold dilutions of virus with a fixed dilution of antiserum in a standard microtitre assay.

Table 2 Reactivity of chimaeric virus with monoclonal antibodies in antigen-blocking and neutralization tests

	Number	Monoclonal antibody Against site	Virus					
			S1/3.10		Sabin type 1		3.370	
			AB	N	AB	N	AB	N
Type 1	955	1	<	<	4	> 3.7	<	<
	237	2	3	ND	3	ND	<	ND
	423	3	4	ND	4	ND	<	ND
Type 3	204	1	2	2.3	<	<	2	2.1
	472	1	3	2.7	<	<	3	2.7
	495	1	4	1.9	<	<	5	> 3.7
	520	1	4	<	<	<	4	<
	175	1	<	<	<	<	<	<
	194	1	<	<	<	<	<	<
	197	1	<	<	<	<	<	<
	199	1	<	<	<	<	<	<
	165	1	<	<	<	<	<	<
	208	1	<	ND	<	ND	<	ND
	880	2	<	ND	<	ND	5	ND
	879	3	<	ND	<	ND	4	ND

Monoclonal antibodies were prepared as described previously¹⁷ and the antigenic sites they recognize identified by testing with a panel of known poliovirus antigenic mutants⁵. Reaction was assessed using the modified single-radial diffusion assay or antigen-blocking test (AB)¹⁷ and neutralization (N). AB titres are given as $\log_{10}$ of the end point effective dilution which inhibits the diffusion of virus. Neutralization titres were determined by challenge with 100 tissue-culture infective doses (TCID₅₀) in a standard microtitre assay and are expressed as $\log_{10}$ of dilution giving 50% inhibition of CPE. ND, not determined; <, less than 1 $\log_{10}$.

suggest that it may be possible to engineer new type 2 and type 3 vaccine strains of comparable safety to the type 1 strain. Our approach to this is based on site-directed mutagenesis of infectious cloned poliovirus complementary DNA of Sabin type 1 with a view to modifying the antigenic determinants of this virus so that they resemble those of poliovirus types 2 and 3.

The major antigenic sites involved in the neutralization of poliovirus type 3 have been identified as residues 89-100 in VP1 (designated site 1), residues 164-172 in VP2 (site 2) and residues 58-60 in VP3 with residues 286-290 in VP1 (sites 3b and 3a respectively)⁴. A 1,174-base pair (bp) *Pst*I restriction endonuclease fragment of an infectious cDNA clone of the Sabin type 1 vaccine strain P1/LS-c, 2ab, which includes the region encoding antigenic site 1, was subcloned into phage M13mp18. Oligonucleotide-directed mutagenesis¹³ was carried out using a 69-base oligonucleotide which resulted in the coding region for eight amino acids (SASTKNKD) from site 1 being replaced by the corresponding amino acids (EQPTTRVQ) from poliovirus type 3 strain 3.370. Strain 3.370 is derived from Sabin type 3 vaccine virus and differs from it only at residue 99 where 3.370 has valine in place of alanine¹⁴. A full length cDNA of the Sabin type 1 strain incorporating the mutated fragment was constructed and used to transfect Hep2-c cells¹⁵. After 9 days a cytopathic effect (CPE) was observed and a virus, designated S1/3.10, which grew more slowly and to lower titre than either the Sabin type 1 or type 3 vaccine strains, was recovered. Sequence analysis of ~300 bases from the appropriate region of virus RNA³ confirmed that this virus was a type 1 poliovirus carrying the expected sequence modification at nucleotides 2,762-2,785 encoding amino acids 93-100 in VP1.

In a standard typing assay the unknown poliovirus is incubated with pairs of monospecific antisera (i.e. type 1 + type 2, type 2 + type 3, type 1 + type 3) to neutralize all but one serotype of virus in each test¹⁶. The chimaera S1/3.10 was neutralized by all three combinations of antisera, by the type 1 or type 3 antisera alone, but not by the type 2 antiserum (Table 1). This demonstrates that the virus chimaera has antigenic characteristics of both type 1 and type 3, but is not a mixture of these

Table 3 Antibody titres against intact and trypsin cleaved virus in antigen-blocking tests

Animal	Immunized with	S1/3.10		Antibody titre to virus		3.370	
		S1/3.10	S1/3.10 TRP	Sabin type 1	Sabin type 1 TRP	3.370	3.370 TRP
Mouse							
1	Sabin type 1	160	160	320	160	<10	<10
2		160	160	160	160	<10	<10
3		640	640	640	640	<10	<10
4		640	640	640	320	<10	<10
5		160	160	160	160	<10	<10
6	S1/3.10	160	160	160	160	10	<10
7		2,560	2,560	2,560	2,560	1,280	<10
8		80	80	80	80	<10	<10
9		80	80	80	80	20	<10
10		160	160	160	160	20	<10
Rabbit							
1	S1/3.10	80	80	80	80	10	<10
2		320	320	80	80	80	<10
3		160	160	40	40	160	<10
Monkey							
1	S1/3.10	160	160	160	160	20	<10

Animals were immunized with 0.1 ml of sucrose purified poliovirus of titre $\sim 10^8$ TCID₅₀ ml⁻¹. Mice were inoculated twice by the intraperitoneal route, and rabbits twice by the intramuscular route. A cynomolgus monkey was fed 1 ml of virus tissue culture fluid and blood samples and faecal specimens taken twice a week for four weeks. Antibody titres were measured in the antigen blocking test¹⁷ and are expressed as the end-point dilution which inhibits the diffusion of virus. Viruses were treated with trypsin as described⁵. TRP, trypsin treated virus.

viruses. The reactivity of S1/3.10 with panels of Sabin 1-specific and Sabin 3-specific monoclonal antibodies^{4,5} was examined using an antigen-blocking test¹⁷ and neutralization (Table 2). All type 3 site 1 antibodies which reacted with type 3 strain 3.370 also reacted with the chimaeric virus, whereas antibodies which reacted with Sabin 3 but not 3.370 also failed to react with S1/3.10. Thus the antigenic activity of the type 3 sequence was faithfully expressed in the chimaeric virus. In addition, all antibodies, such as 955, which reacted with site 1 of Sabin 1 failed to react with the chimaera confirming that this site had been lost. As expected, monoclonal antibodies against other type 1 sites (for example 237, 423) reacted with the chimaeric virus whereas antibodies to other type 3 sites (for example 879, 880) did not. The antibodies gave parallel results in neutralization tests.

The immunogenic potential of the chimaera S1/3.10 was examined in mice and rabbits (Table 3). Mice immunized with Sabin type 1 produced antibody against both S1/3.10 and Sabin type 1, but produced no detectable antibody to the type 3 strain 3.370. In contrast both mice and rabbits immunized with S1/3.10 produced significant levels of antibody reacting with both the type 3 strain 3.370 and Sabin type 1.

The specificity of the type 3 antibody induced by the chimaera was examined further. Antigenic site 1 of polioviruses types 1 and 3 can be specifically cleaved by trypsin, thus antibodies directed against this site fail to react with cleaved virus^{5,18}. The antisera to S1/3.10 were therefore tested against trypsin-cleaved virus to confirm that the type 3-specific component was directed against the type 3 site 1 of S1/3.10 rather than possible cross-reactive sites. These antisera failed to react with trypsin-cleaved type 3 indicating that the type 3 response was indeed directed against the intact site 1 in the configuration found in normal infectious virus.

It has been shown that antigenic site 1 of virus excreted by vaccinees is cleaved¹⁹, presumably by intestinal proteases. This implies that virus in the gut presents a different antigenic configuration from that grown in tissue culture, and presumably also from virus produced in other sites of the body. The configuration of virus which induces the humoral immune response during a natural infection is unknown. A cynomolgus monkey was therefore fed the chimaera S1/3.10 and faecal and blood samples collected over a period of four weeks. Virus was isolated from the faeces for several days, indicating that a gut infection

had been established. At 14 days antibody against both S1/3.10 and Sabin type 1 were detected in the serum (Table 3). There was also detectable antibody against intact, but not trypsin cleaved, type 3 virus 3.370. These findings indicate that the chimaera induced antibodies against its engineered type 3 sequence following a gut infection, and suggest that at least a proportion of the response was induced by virus generated outside the intestine.

In summary, we have shown that after exchange of one of its antigenic sites the very safe poliovirus type 1 Sabin vaccine strain can induce neutralizing antibodies against poliovirus type 3. The modification involves a region of the genome which is unlikely to contain attenuating mutations in this strain¹⁰ and therefore the chimaera should retain the stable attenuation phenotype. A comprehensive collection of type 1/3 and type 1/2 chimaeras involving all three antigenic sites is now under construction. These viruses should provide valuable tools to study further both cellular and humoral immunity to poliomyelitis. Appropriate modifications of the poliovirus type 1 Sabin strain using these techniques could provide an approach not only to the development of new improved type 2 and type 3 polio vaccines but also to vaccines against other enteroviruses, such as the hepatitis A virus.

Received 17 November 1987, accepted 13 January 1988.

- Hogle, J. M., Chow, M. & Filman, D. J. *Science* **229**, 1358-1365 (1985).
- Minor, P. D. *et al. Nature* **301**, 674-679 (1983).
- Minor, P. D. *et al. J. gen. Virol.* **66**, 1159-1165 (1985).
- Minor, P. D., Ferguson, M., Evans, D. M. A., Almond, J. W. & Icenogle, J. P. *J. gen. Virol.* **67**, 1283-1291 (1986).
- Minor, P. D. *et al. J. gen. Virol.* **68**, 1857-1865 (1987).
- Diamond, D. *et al. Science* **229**, 1090-1093 (1985).
- Blondel, B. *et al. J. Virol.* **57**, 81-90 (1986).
- Sabin, A. B., Boulger, L. R. *J. biol. Stand.* **1**, 115-118 (1973).
- WHO Collaborative Study *J. biol. Stand.* **9**, 163-184 (1981).
- Omata, T. *et al. J. Virol.* **58**, 348-358 (1986).
- Westrop, G. D. *et al. in The molecular biology of positive strand viruses* (eds Rowlands, D. J., Mahy, B. W. J. & Mayo M.) 53-60 (Academic, London, 1986).
- Westrop, G. D. *et al.* (in preparation).
- Kramer, W., Drutsa, V., Jansen, H., Kramer, B., Pflugfelder, M. & Fritz, H. *Nucleic Acids Res.* **12**, 9441-9456 (1984).
- Evans, D. M. A., Minor, P. D., Schild, G. C. & Almond, J. W. *Nature* **304**, 459-462 (1983).
- Parker, B. A. & Stark, G. R. *J. Virol.* **31**, 360-375 (1979).
- Domok, I. & Magrath, D. I. *WHO Offset Publication No.* **46** (1979).
- Ferguson, M. *et al. J. gen. Virol.* **65**, 197-201 (1984).
- Icenogle, J. P., Minor, P. D., Ferguson, M. & Hogle, J. M. *J. Virol.* **60**, 297-301 (1986).
- Minor, P. D., John, A., Ferguson, M. & Icenogle, J. P. *J. gen. Virol.* **67**, 693-706 (1987).

Autocrine generation and requirement of BSF-2/IL-6 for human multiple myelomas

Michio Kawano, Toshio Hirano*, Tadashi Matsuda*, Tetsuya Taga*, Yasuhiro Horii*, Koji Iwato, Hideki Asaoku, Bo Tang*, Osamu Tanabe, Hideo Tanaka, Atsushi Kuramoto & Tadimitsu Kishimoto†

Department of Internal Medicine, Research Institute for Nuclear Medicine and Biology, Hiroshima University, Kasumi 1-2-3, Minami-ku, Hiroshima 735, Japan

* Division of Immunology, Institute for Molecular and Cellular Biology, Osaka University, 1-3 Yamada-Oka, Suita 565 Japan

Human B cell stimulatory factor 2 (BSF-2) was originally characterized and isolated as a T cell-derived factor that caused the terminal maturation of activated B cells to immunoglobulin-producing cells^{1,2}. Molecular cloning of the complementary DNA predicts that BSF-2 is a protein of relative molecular mass (M_r) 26,000 similar or identical to interferon β_2 , hybridoma plasmacytoma growth factor and hepatocyte stimulating factor³⁻⁷. IL-6 has been proposed as a name for this molecule^{8,9}. It is now known that BSF-2 has a wide variety of biological functions and that its target cells are not restricted to normal B cells^{1,10}. Responses are also seen in T cells^{11,12}, plasmacytomas⁶, hepatocytes^{7,13}, haematopoietic stem cells¹⁴, fibroblasts¹⁵ and rat pheochromocytoma, PC12 (Satoh, T. *et al.*, manuscript in preparation). Of particular interest to this report is that human BSF-2 is a potent growth factor for murine plasmacytomas and hybridomas^{6,8,10}. This observation suggested to us that constitutive expression of BSF-2 or its receptor could be responsible for the generation of human myelomas. In this study we report that myeloma cells freshly isolated from patients produce BSF-2 and express its receptors. Moreover, anti-BSF-2 antibody inhibits the *in vitro* growth of myeloma cells. This is direct evidence that an autocrine loop is operating in oncogenesis of human myelomas.

Myeloma cells were obtained prior to chemotherapy from 26 patients diagnosed as having multiple myeloma of IgG type and in the advanced stage III¹⁶. Recombinant BSF-2 (rBSF-2) as well as a partially purified myeloma cell growth factor (MCGF)¹⁷ augmented the proliferation of purified myeloma cells by about two to 20-fold in 12 out of 26 patients. Representative data on two different responder myelomas are shown in Fig. 1. As shown, rBSF-2 augmented ³H-thymidine (³H-TdR) uptake in myeloma cells in a dose-dependent manner. Furthermore, the activity of partially purified MCGF was neutralized by anti-BSF-2 antibodies as shown in Fig. 1b. The data indicate that BSF-2 is in fact the growth factor for human myeloma cells and that the previously reported MCGF is identical with BSF-2.

We next examined whether the responsiveness of myeloma cells to rBSF-2 is related to the expression of BSF-2 receptor (R). Scatchard analysis of six different myelomas (three were responders and the others were non-responders to exogenously added BSF-2) gave dissociation constants ranging from 1.5 to 7.9×10^{-10} M and estimated the numbers of BSF-2R per cell at between 140 and 890. A representative Scatchard plot of responder myeloma cells is shown in Fig. 2. The data indicate that all the myelomas express BSF-2R but that the number of receptors does not correlate with responsiveness to exogenously added BSF-2 (figure legend for Fig. 2).

The possibility that BSF-2 is produced by the myeloma cells was then investigated. As shown in Fig. 3a, the culture supernatant of freshly isolated myeloma cells augmented ³H-TdR

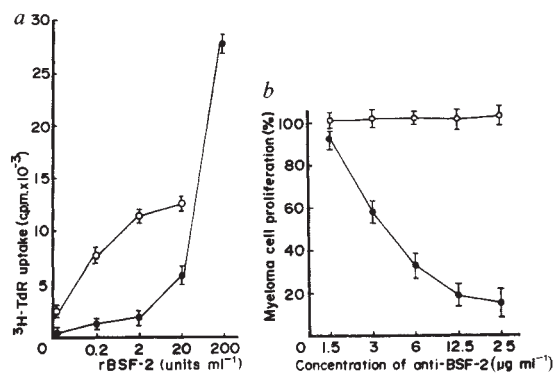


Fig. 1 Recombinant BSF-2-induced ³H-TdR uptake in freshly isolated myeloma cells and the neutralization of the activity of partially purified MCGF by the BSF-2-specific antibodies. *a*, Myeloma cells from case 8 (○) and 26 (●) were cultured with various concentrations of rBSF-2. The results represent the mean $\pm$ s.e. of triplicate cultures. *b*, Myeloma cells from case 8 were cultured for 48 h with 10% v/v of partially purified MCGF in the presence of various concentrations of IgG-fraction of pre-immune rabbit sera (○) or IgG-fraction of rabbit antibodies specific to BSF-2 (●)²⁵. The ³H-TdR uptakes in the presence or absence of MCGF were 12,600 c.p.m. and 3,150 c.p.m., respectively. The data indicate the specific increase in ³H-TdR uptake in the experimental groups as a percentage of that induced by MCGF alone. The specific increase in ³H-TdR uptake was calculated by subtracting the background (3,150 c.p.m.) from the experimental groups.

Methods. Freshly isolated myeloma cells (more than 95% pure¹⁷) were cultured for 48 h in 0.2 ml of RPMI-1640 medium (Nissui, Japan) supplemented with 10% fetal calf serum (FCS, Microbiological Associates) and 1×10^{-5} M 2-mercaptoethanol at 4×10^6 cells per well in 96-well plates (Corning) in the presence of various concentrations of rBSF-2 (ref. 25). The cells were pulsed with 1 μ Ci per well of ³H-TdR (5 Ci mmol⁻¹, New England Nuclear) for the last 16 h of incubation. MCGF was partially purified from the concanavalin A-stimulated peripheral blood mononuclear cells (PBMNC) as described previously¹⁷.

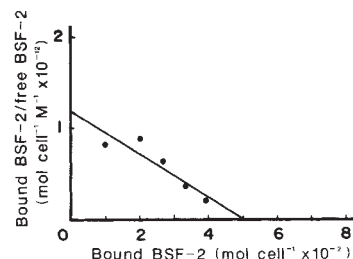


Fig. 2 Expression of BSF-2R in freshly isolated myeloma cells. Scatchard plot analysis of BSF-2R was performed as described previously utilizing ¹²⁵I-BSF-2 (specific activity was 5.5×10^{12} c.p.m. g⁻¹)¹⁸. The data on a representative case (8) is illustrated. The numbers of receptor per cell and K_d values of 6 different myelomas are as follows; case 8, 500, 4.5×10^{-10} M; case 21, 180, 1.5×10^{-10} M; case 26, 890, 2.2×10^{-10} M; case 10, 140, 4.8×10^{-10} M; case 13, 310, 5.5×10^{-10} M; case 17, 390, 7.9×10^{-10} M. Cases 8, 21 and 26 were responders to BSF-2 and the others were non-responders.

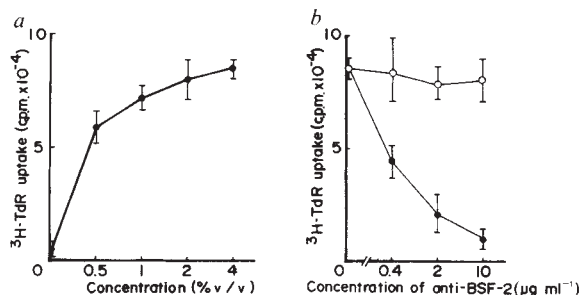
uptake of a murine BSF-2-dependent hybridoma clone, MH60.BSF2, in a dose-dependent manner. Furthermore, the growth activity of the culture supernatant was neutralized by antibodies specific to BSF-2 (Fig. 3b), indicating that myeloma cells generate BSF-2. The culture supernatants of eight additional myelomas (five rBSF-2 responders and three non-responders) and a myeloma cell line, U266, which was previously shown to express BSF-2R¹⁸, were also found to contain BSF-2 (0.44 to 6.9 units ml⁻¹). Moreover, BSF-2 mRNA was detected in all four freshly isolated myeloma cell cultures examined (one responder and three non-responders) as well as in U266

† To whom correspondence should be addressed.

Fig. 3 Constitutive generation of BSF-2 by myeloma cells. *a*, BSF-2 activity of the culture supernatant of myeloma cells from a representative case (8). *b*, Neutralization of the BSF-2 activity of the culture supernatant of myeloma cells by anti-BSF-2 antibodies. MH60.BSF2 cells were cultured with various concentrations of the IgG-fraction of anti-BSF-2 antibodies (●) or the IgG-fraction of pre-immune rabbit serum (○) in the presence of the culture supernatant of myeloma cells (case 8) at the concentration of 4% (v/v). *c*, Expression of BSF-2 mRNA in myeloma cells. Total cytoplasmic RNA from freshly isolated PBMNC (row 1), PBMNC stimulated with phytohaemagglutinin ($1 \mu\text{g ml}^{-1}$) and 12-*o*-tetradecanoyl-phorbol-14 acetate (5 ng ml^{-1})

for 20 h (row 2), freshly isolated myeloma cells (rows 3–6) representing cases 8, 10, 13 and 17, respectively and U266 cell (row 7) was hybridized with a ^{32}P -labelled complementary DNA probe (*TaqI/BanII* fragment of pBSF2.38)³. The quantity of RNA is halved in each successive dot; the first dot line (7) contains RNA from 2×10^6 cells in rows 3–7 and 2.5 μg in rows 1 and 2.

Methods. Myeloma cells were cultured ($1 \times 10^6 \text{ ml}^{-1}$ per well) in 24-well microplate for two days and the culture supernatants were measured for BSF-2 activity utilizing a BSF-2 dependent murine hybridoma clone, MH60.BSF2 (T.M. *et al.*, manuscript submitted). MH60.BSF2 cells were cultured in 96-well microplates at 10^4 cells per well with various concentrations of the culture supernatants of myeloma cells and pulsed with ^3H -TdR ($1 \mu\text{Ci}$ per cell) during the last 6 h of the total 48 h culture period. Total cytoplasmic RNA was extracted using a modification of the method described by Chirgwin *et al.*²⁶ and blotted onto nylon membrane (GeneScreen Plus, NEN Research Products).



(Fig. 3c). The data indicate that myeloma cells constitutively produce BSF-2 *in vitro* and *in vivo* regardless of their responsiveness to rBSF-2.

To show directly that BSF-2 functions as an autocrine growth factor for the generation of myelomas, we examined whether BSF-2 specific antibodies could inhibit the *in vitro* growth of myeloma cells. The anti-BSF-2 antibodies inhibited the spontaneous ^3H -TdR uptake of four different responder myelomas (Fig. 4a shows data on a representative case reducing ^3H -TdR incorporation by between 40 and 70% at a concentration of $50 \mu\text{g ml}^{-1}$). The inhibitory effect could be overcome by the addition of rBSF-2 at a concentration of 20 units per ml, indicating that it was indeed specific to BSF-2. However, it should be noted that the anti-BSF-2 antibodies had no inhibitory effect on five non-responder myelomas that were examined.

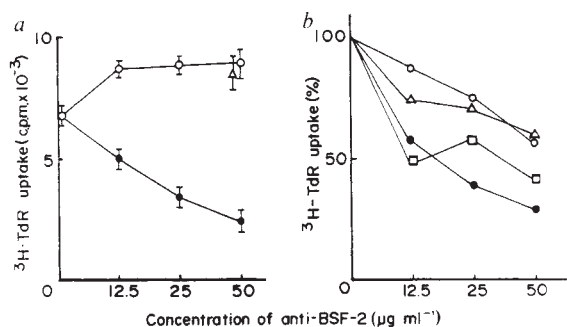


Fig. 4 Inhibition of ^3H -TdR uptake of freshly isolated responder myeloma cells by antibodies specific to BSF-2. *a*, Freshly isolated responder myeloma cells (case 21) were cultured at 2×10^4 per well in the presence of various concentrations of the IgG-fraction of anti-BSF-2 antibodies (●) or that of pre-immune rabbit serum (○) for two days and the ^3H -TdR uptake was measured as described in the legend for Fig. 1. The ^3H -TdR uptake of myeloma cells in the presence of both anti-BSF-2 ($50 \mu\text{g ml}^{-1}$) and rBSF-2 (20 units ml^{-1}) is also shown (△). The results represent the mean $\pm$ S.E. of triplicate cultures. *b*, ^3H -TdR uptake of four different myelomas in the presence of various concentrations of anti-BSF-2 as a percentage of the ^3H -TdR uptake in the presence of the equivalent amounts of normal rabbit IgG. (○) (●) (□) and (△) represent cases 8, 21, 22 and 23. The spontaneous ^3H -TdR uptakes of cells from cases 8, 21, 22, and 23 were 5,390, 6,730, 550 and 3,030 c.p.m., respectively. The BSF-2 activities of the culture supernatants of cases 8, 21, 22 and 23, determined as described in the legend of Fig. 3 were 4.4, 0.96, 0.84 and 3.3 units ml^{-1} , respectively.

In this report, we demonstrate that freshly isolated myeloma cells constitutively produce BSF-2 and express BSF-2R. Furthermore, certain myeloma cells proliferate in response to exogenous BSF-2 and anti-BSF-2 antibodies inhibit ^3H -TdR uptake by those responder myeloma cells. Taken together the data indicate that BSF-2 is functioning as an autocrine growth factor for human myelomas.

It is well known that the intraperitoneal injection of mineral oil generates plasmacytomas in mice¹⁹ and that it enhances the expansion of hybridoma cells. Interestingly, it has been shown that adherent peritoneal cells can be induced by mineral oil to produce plasmacytoma growth factor²⁰, which could be the murine equivalent of human BSF-2 in 50-fold greater amounts. These data suggest that BSF-2 could also be important for the generation and/or clonal expansion of myeloma cells *in vivo*.

There has been accumulating circumstantial evidence to support the hypothesis, originally proposed by Sporn and Todaro, that autocrine stimulation is a mechanism of transformation^{21,22}. For instance, antibodies to growth factors such as PDGF or bombesin have been shown to inhibit the *in vivo* or *in vitro* growth of several established cell lines^{23,24}. In this report, we provide the first evidence that the growth of myeloma cells freshly isolated from patients can be inhibited by anti-BSF-2 antibody. Specific inhibitors of BSF-2 are therefore potential candidates for use in the therapy of human myelomas.

This work was supported in part by grants from the Ministry of Education, Science and Culture. We thank Drs Edward Barsumian and Tadatsuga Taniguchi for their critical review of the manuscript, and Ms. K. Kubota, J. Shibuya, M. Kawata, Y. Ohto and H. Sumida for their secretarial and technical assistance.

Received 4 November 1987; accepted 8 January 1988.

- Hirano, T. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 5490–5494 (1985).
- Hirano, T. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 228–231 (1987).
- Hirano, T. *et al.* *Nature* **324**, 73–76 (1986).
- Haegeman, G., Content, J., Volckaert, G., Derynck, R., Tavernier, J. & Fiers, W. *Eur. J. Biochem.* **159**, 625–632 (1986).
- Zilberstein, A., Ruggieri, R., Korn, J. H. & Revel, M. *EMBO J.* **5**, 2529–2537 (1986).
- Van Damme, J. *et al.* *J. exp. Med.* **165**, 914–919 (1987).
- Andus, T. *et al.* *FEBS Lett.* **221**, 18–22 (1987).
- Poupart, P. *et al.* *EMBO J.* **6**, 1219–1224 (1987).
- Yasukawa, K. *et al.* *EMBO J.* **6**, 2939–2945 (1987).
- Muraguchi, A. *et al.* *J. exp. Med.* (in the press).
- Garman, R. D., Jacobs, K. A., Clark, S. C. & Raulet, D. H. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7629–7633 (1987).
- Lotz, M. *et al.* *J. exp. Med.* (in the press).
- Gauldie, J., Richards, C., Harnish, D., Lansdorp, P. & Baumann, H. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7251–7255 (1987).

14. Ikebuchi, K., Wong, G. G., Clark, S. C., Ihle, J. N., Hirai, Y. & Ogawa, M. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
15. Kohase, M., Henriksen-DeStefano, D., May, L. T., Vilcek, J. & Sehgal, P. B. *Cell* **45**, 659-666 (1986).
16. Durie, B. G. & Salmon, S. E. *Cancer* **36**, 842-854 (1975).
17. Kawano, M., Fujimura, K., Iwato, K., Asaoku, H., Tanabe, O. & Kuramoto, A. *Acta Haemat. jap.* **49**, 1740-1748 (1986).
18. Taga, T., Kawanishi, K., Hardy, R. R., Hirano, T. & Kishimoto, T. *J. exp. Med.* **166**, 967-981 (1987).
19. Potter, M. & Boyce, C. *Nature* **193**, 1086-1087 (1962).
20. Nordan, R. P. & Potter, M. *Science* **233**, 566-569 (1986).
21. Sporn, M. B. & Todaro, G. J. *New Engl. J. Med.* **303**, 878-880 (1980).
22. Sporn, M. B. & Roberts, A. B. *Nature* **313**, 745-747 (1985).
23. Huang, J. S., Huang, S. S. & Deuel, T. F. *Cell* **39**, 79-87 (1984).
24. Cuttitta, F. *et al.* *Nature* **316**, 823-826 (1986).
25. Hirano, T., Matsuda, T., Hosoi, K., Okano, A., Matsui, H. & Kishimoto, T. *Immun. Lett.* **17**, 41-45 (1988).
26. Chirwin, J. M., Przybla, A. E., MacDonald, R. L. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).

Multiple endocrine neoplasia type 1 gene maps to chromosome 11 and is lost in insulinoma

Catharina Larsson*, Britt Skogseid†, Kjell Öberg‡, Yusuke Nakamura§ & Magnus Nordenskjöld*

* Department of Clinical Genetics, Karolinska Hospital, S-10401 Stockholm, Sweden

† Department of Internal Medicine and ‡ Ludwig Institute for Cancer Research, Uppsala Branch, University Hospital, S-75185 Uppsala, Sweden

§ Howard Hughes Medical Institute, University of Utah Medical Center, Salt Lake City, Utah 84132, USA

Multiple endocrine neoplasia type 1 (MEN-1) is a predisposition to hyperplasia of the parathyroid glands, and to hyperplasia or tumours of the anterior pituitary and the endocrine pancreas, and is inherited as an autosomal dominant trait¹. Here we map the *MEN-1* locus to chromosome 11 by family studies, and demonstrate tight linkage with the human muscle phosphorylase gene. By comparing constitutional and tumour tissue genotypes of insulinomas from a pair of brothers who had inherited MEN-1 from their mother, we have shown that oncogenesis in these cases involves unmasking of a recessive mutation at this locus.

Gene deletions have been implicated as the ultimate event in oncogenesis of inherited neoplasia^{2,3}. According to the two-mutation model, an inherited tumour results from unmasking of a recessive mutation at the disease locus. The first mutation is carried in the germ-line and predisposes to neoplasia, whereas a second, somatic event serves to eliminate the remaining wild-type allele at this locus. This mechanism has been confirmed for retinoblastoma⁴⁻⁷ and suggested for bilateral acoustic neuroma^{8,9} and for colon carcinoma associated with adenomatous polyposis of the colon^{10,11}. In these tumours, the second mutational event is often a chromosome mutation, which eliminates a large region of one chromosome.

Based on the hypothesis that oncogenesis in MEN-1 results from two mutational events affecting the *MEN-1* locus, we postulated that chromosomal rearrangements in *MEN-1*-associated tumours could indicate the chromosome on which the *MEN-1* locus is found. We therefore compared the constitutional and tumour genotypes at different loci in two brothers with insulinoma who had inherited MEN-1 from their mother (Fig. 1a). These studies took advantage of polymorphic restriction enzyme recognition sites at different loci defined by recombinant DNA probes.

As shown in Table 1, the constitutional genotype was retained in tumour tissue at informative loci on 17 different chromosomes in these tumours. But both tumours showed loss of one constitutional allele at all informative loci on chromosome 11 (Fig. 1, Table 1). Cases II-1 and II-3 were heterozygous at 6 and 8 loci on chromosome 11 respectively, spanning from band p15

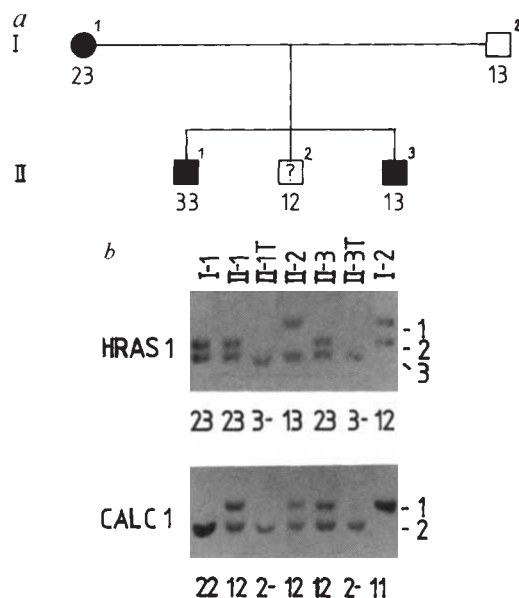


Fig. 1 a, Pedigree of the LU kindred with genotypes for the *PYGM* locus (probe pMCMPI) indicated below each individual. II-2 showed no signs of MEN-1 at the age of 33, but was scored as unknown (?) for the *MEN-1* phenotype because a later onset of the disease cannot be excluded^{1,14}. b, Autoradiograms showing the genotypes for the LU kindred at the *INS* and *CALC-1* loci¹². Alleles are indicated by numbers to right. The genotypes for each individual (indicated above) are given below the autoradiograms. T indicates the lane with tumour tissue DNA. Experimental procedures as previously described¹³.

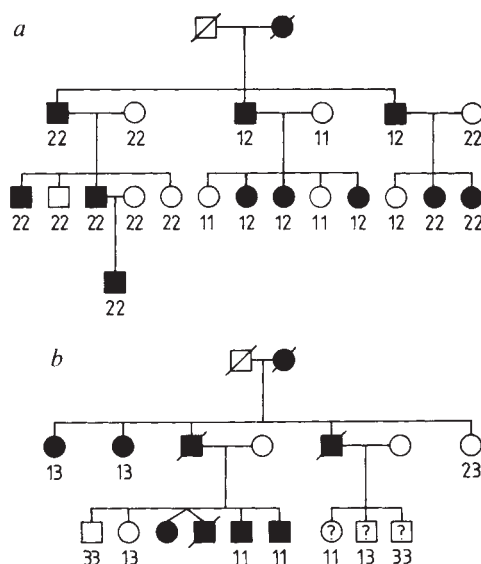


Fig. 2 Pedigrees of a, KA and b, EZ kindreds with the genotype at the *PYGM* locus indicated below each individual. Individuals denoted by a question mark were scored as unknown for the *MEN-1* locus as they had not yet shown signs of the disease according to established diagnostic criteria, but had not passed the age of 35 years^{1,14}.

(*HRAS-1*) to band q23 (*ApoA-1*) (ref. 12). One allele was lost in the tumour tissues at each informative locus on this chromosome. As displayed in Fig. 1b, the rearrangement eliminated the paternally derived alleles at the *HRAS-1* and *CALC-1* loci. Similar results were obtained for all other loci studied on chromosome 11 (Table 1). The allelic copy number on chromosome 11 in normal and tumour tissue was determined by quantitative densitometry. A 50% reduction of the autoradiograph

Table 1 Genotypes of cases II-1 and II-3 of the LU kindred in normal tissue (N) and tumour tissue (T) DNA at different chromosomal loci

Chromosome 11										
Locus	HRAS1	INS	CALC1	PTH	CAT	PGA	PYGM	D11S84	APOA1	
Enzyme	(<i>TaqI</i>)	(<i>TaqI</i>)	(<i>TaqI</i>)	(<i>PstI</i>)	(<i>TaqI</i>)	(<i>EcoRI</i>)	(<i>TaqI</i>)	(<i>TaqI</i>)	(<i>PstI</i>)	
Case no.	N T	N T	N T	N T	N T	N T	N T	N T	N T	
II-1 LU	2, 3 3, -	1, 2 1, -	1, 2 2, -	1, 2 2, -	1, 1	1, 2 1, -	3, 3	2, 2	1, 2 1, -	
II-3 LU	2, 3 3, -	1, 2 1, -	1, 2 2, -	1, 2 2, -	1, 2 1, -	1, 3 1, -	1, 3 3, -	1, 2 1, -	2, 2	
Chromosome 11										
Locus	APOA2	D2S6	D3S2	D5S2	D6S10	MetH	CA-2	IFNA	D10S4	
Enzyme	(<i>MspI</i>)	(<i>TaqI</i>)	(<i>MspI</i>)	(<i>MspI</i>)	(<i>TaqI</i>)	(<i>MspI</i>)	(<i>TaqI</i>)	(<i>MspI</i>)	(<i>TaqI</i>)	
Case no.	N T	N T	N T	N T	N T	N T	N T	N T	N T	
II-1 LU	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	1, 1	1, 2, 3 1, 2, 3	2, 3 2, 3	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	
II-3 LU	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	2, 3 2, 3	1, 2 1, 2	1, 2 1, 2	1, 1	
Chromosome 11										
Locus	D12S7	D13S1	D15S1	D17S1	APOC2	D2OS4	MB	DXYS1		
Enzyme	(<i>TaqI</i>)	(<i>MspI</i>)	(<i>MspI</i>)	(<i>MspI</i>)	(<i>TaqI</i>)	(<i>MspI</i>)	(<i>TaqI</i>)	(<i>TaqI</i>)		
Case no.	N T	N T	N T	N T	N T	N T	N T	N T		
II-1 LU	2, 3 2, 3	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	X, Y X, Y		
II-3 LU	2, 3 2, 3	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	1, 2 1, 2	X, Y X, Y		

DNA samples from normal tissue (N) and tumour tissue (T) of insulinomas from cases II-1 and II-3 in the LU kindred were digested with the restriction endonucleases indicated, separated by electrophoresis and transferred to nylon membranes. The Southern blots were hybridized to radiolabelled DNA probes¹², washed and autoradiographed as described¹³. Maternal alleles shown in bold.

signals at all loci studied was found, indicating loss of one chromosome complement in the tumour tissue (data not shown). This excludes the possibility that tumorigenesis in these insulinomas involved mitotic recombination with retained chromosome number^{4,15}.

DNA from hyperplastic pancreatic tissue of case II-3 in the LU kindred was also studied. This tissue was classified histopathologically as benign β -cell hyperplasia and showed the same type of chromosome 11 rearrangements (Fig. 1b) as the malignant insulinomas described above (data not shown).

To confirm the tentative localization of *MEN-1* locus to chromosome 11, the linkage of *MEN-1* to different restriction fragment length polymorphism markers was analysed in four families, including the LU kindred. In these families, *MEN-1* was unlinked to *INS* (regionally localized to 11p15 (ref. 12) and *ApoA-1* (at 11q23). But when markers localized closer to the centromere of chromosome 11 were analysed, we found evidence in favour of linkage. The data were analysed with the computer program LIPED (ref. 16) to calculate lod scores, assuming various recombination fractions and full penetrance for *MEN-1* at the age of 35 years. Individuals above this age, who had been subject to at least three extensive biochemical screenings during a ten-year follow-up period and did not show signs of *MEN-1*-associated lesions¹⁴, were scored as unaffected. In three informative families, the LU (Fig. 1a), the KA (Fig. 2a) and the EZ kindred (Fig. 2b), *MEN-1* was found to be closely linked to the skeletal muscle glycogen phosphorylase (*PYGM*) locus. No meiotic recombinants were detected for this marker, giving a maximal lod score of 4.37 at a recombination distance of 0% between *MEN-1* and *PYGM*. The *PYGM* locus corresponds to the McArdle's disease (muscle phosphorylase deficiency) locus and has been regionally assigned to chromosome 11q by mapping to a somatic-cell hybrid panel¹⁷. Multipoint linkage studies have mapped *PYGM* between *PGA*,

which is within the pericentric region of chromosome 11, and *D11S84* (ref. 18). Based on linkage data, the *MEN-1* locus is clearly located in the immediate vicinity of *PYGM* within the centromeric region on the long arm of chromosome 11.

The mapping of *MEN-1* to chromosome 11 implies that the loss of alleles in the tumour tissue of case II-1 and II-3 (Fig. 1b and Table 1) must have eliminated the wild-type allele at this locus. Not only was the paternally derived allele of the closely linked *PYGM* affected in this manner, but also flanking markers unlinked to *MEN-1* and scattered along the whole chromosome 11 from *HRAS* at p15 to *ApoA-1* at q23. Thus oncogenesis in *MEN-1* involves unmasking of a recessive mutation at the *MEN-1* locus, in agreement with the two-mutant model^{2,3}. Furthermore, the localization of the *MEN-1* locus to the long arm of chromosome 11 implies that a new recessive cancer gene has been discovered. The localization of this gene differs from the previously known genes on the short arm of chromosome 11 implicated in tumorigenesis of Wilms' tumour, rhabdomyosarcoma and hepatoblastoma and located within bands 11p13 to 11p15 (refs 3, 15, 19).

These data suggest that the wild-type allele of the *MEN-1* locus regulates the normal differentiation of the anterior pituitary, parathyroid glands, and endocrine pancreas, namely the tissues at risk for neoplastic transformation in *MEN-1* (ref. 3). The *MEN-1* predisposition would be a constitutional mutation in heterozygous form, inherited as an autosomal dominant trait. Tumour development involves a second mutational event, which in two tumours studied involved the chromosome 11 carrying the remaining wild-type allele at the *MEN-1* locus by means of a chromosome-loss event (Fig. 1). It should be pointed out that it is not yet known if lesions in *MEN-1* other than insulinoma and β -cell hyperplasia show similar chromosomal rearrangements, and if sporadic tumours of the same type also result from similar events.

Table 2 Lod score for linkage of *MEN-1* to *PYGM*

		Recombination fraction (θ)							
Marker locus	Family	0.00	0.01	0.05	0.10	0.20	0.30	0.40	
PYGM	LU	0.30	0.30	0.26	0.21	0.13	0.06	0.02	
	KA	2.26	2.26	2.05	1.83	1.35	0.84	0.31	
	EZ	1.81	1.80	1.63	1.44	1.03	0.59	0.17	
	Total	4.37	4.36	3.93	3.48	2.52	1.49	0.50	

Peak lod score ($\hat{z}$) was 4.37 at $\hat{\theta} = 0$.

Finally, we would like to stress that the regional localization of the *MEN-1* locus is based on only three informative families. Therefore it cannot be excluded that additional loci might be associated with the *MEN-1* phenotype.

We thank Drs W. Cavenee, T. Dryja, A. von Gabain, J. Gusella, S. Humphries, J. Höppner, A. Jeffereys, M. Leppart, M. Litt, H. Luthman, H. Mayer, P. O'Connell, L.-C. Tsui, G. Vande Woude, R. White, J. Scott, P. Szabo, B. Zelle and the American Type Culture Collection for providing recombinant DNA probes, J. Attwood for supplying the 'Linksys' computer program and Drs K.-H. Gustavsson, D. Hellgren and L. Iselius for valuable suggestions. This work was supported by grants from the Swedish Cancer Society, King Gustav V's Jubilee Foundation, the Inga Britt and Arne Lundberg Foundation and the Swedish Work Environment Fund.

Received 11 December 1987; accepted 5 January 1988.

1. Yamaguchi, K., Kameya, T. & Abe, K. *Clin. Endocrinol. Metab.* **9**, 261-284 (1980).
2. Knudsen, A. G. Jr *Proc. natn. Acad. Sci. U.S.A.* **68**, 820-823 (1971).
3. Nordenskjöld, M. & Cavenee, W. K. in *Important Advances in Oncology* Vol. IV (ed. De Vita, J. B.) 83-101 (Lippincott, Philadelphia, 1988).
4. Cavenee, W. K. *et al. Nature* **309**, 170-172 (1983).
5. Cavenee, W. K. *et al. Science* **228**, 501-503 (1985).
6. Friend, S. *et al. Nature* **323**, 643-646 (1986).
7. Lee, W.-H. *et al. Nature* **329**, 642-645 (1987).
8. Seizinger, B. R., Martuza, R. L. & Gusella, J. F. *Nature* **322**, 644-647 (1986).
9. Rouleau, G. *et al. Nature* **329**, 246-248 (1987).
10. Bodmer, W. F. *et al. Nature* **328**, 614-616 (1987).
11. Solomon, E. *et al. Nature* **328**, 616-619 (1987).
12. Pearson, P. *et al. Cytogenet. Cell Genet.* (in the press).
13. Lundberg, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 2372-2376 (1987).
14. Skogseid, B. *et al. J. clin. Endocr. Metab.* **64**, 1233-1240 (1987).
15. Scrabble, H. J. *et al. Nature* **329**, 645-647 (1987).
16. Ott, J. *Am. J. hum. Genet.* **28**, 528-529 (1976).
17. Lebo, R. *et al. Science* **225**, 57-59 (1984).
18. Leppert, M. *et al. Cytogenet. Cell Genet.* (in the press).
19. Koufus, A. *et al. Nature* **316**, 330-334 (1985).

Cooperativity of the glucocorticoid receptor and the CACCC-box binding factor

Roland Schüle, Marc Muller, Hidetsugu Otsuka-Murakami* & Rainer Renkawitz

Max-Planck-Institut für Biochemie, Genzentrum, D 8033 Martinsried, FRG

Glucocorticoid receptor binding sites (GRE) are often tightly clustered with other transcription factor binding sequences. Examples of this occur upstream of the genes for chicken lysozyme^{1,2} and human metallothionein II_A (ref. 3), in several retroviral LTRs⁴⁻⁶ and upstream of the rat tryptophan oxygenase (TO) gene⁷. In the TO gene, sequences immediately upstream of a glucocorticoid receptor binding site are required for steroid induction⁷ and contain a CACCC-box identical to that found in the β globin gene. Here we demonstrate specific binding to this TO-CACCC element and show that it will also act cooperatively with a MMTV glucocorticoid receptor binding site. The response to dexamethasone is independent of the order and relative orientation of these elements but does depend on their precise spacing. Optimal induction occurs at a periodicity of approximately 10 base pairs (bp) indicating a requirement for stereospecific alignment. Binding to the CACCC box, however, is not affected by its distance from the glucocorticoid receptor site. We conclude that the observed cooperativity is mediated by protein:protein interactions and does not depend on cooperative DNA binding.

Analysis of the glucocorticoid induction of the TO gene revealed the requirement of additional sequences adjacent to

the receptor binding site⁷. This region (Fig. 1b) contains in antisense orientation a sequence of 11 nucleotides identical to the CACCC-box element⁸ of the β globin promoter. DNase I footprinting experiments of TO sequences (Fig. 1a) demonstrate that the sequence from -472 to -444, which is required for the glucocorticoid response, contains only one protected region (-455 to -439), which harbours the CACCC sequence. Under these conditions the binding of the glucocorticoid receptor cannot be seen (a footprint can only be generated by incubation of purified receptor with DNA in the presence of PEG)⁷. For comparison, the footprint borders of the purified receptor and its close contact sites⁷ are indicated (Fig. 1b). The identity of the CACCC-box binding protein is confirmed by the fact that footprint competition in the -445 region (Fig. 1a) occurs with the double stranded oligonucleotides (Fig. 1b) TO-CACCC(L), CACCC (a shorter oligonucleotide of the TO upstream region, -452 to -438) and β globin (a 22-bp oligonucleotide containing the β globin upstream sequence, -95 to -81), but not with TO-CACCC(L)* or TO-CACCC(S)*, both of which contain mutations which reduce CACCC-box function in the β globin gene⁹. Competition with the MMTV-GRE sequence was also inefficient. In summary, all sequences containing the wildtype β globin CACCC box compete efficiently, whereas the mutants and the MMTV-GRE sequence do not compete.

To distinguish between effects specific to the TO gene and general mechanisms of steroid induction, we combined CACCC-box sequences with the well characterized GRE of MMTV (-186 to -170)^{10,11}. These sequences (Fig. 1b) were cloned into ptkCAT, upstream (-105) of the thymidine kinase (tk) promoter which controls the chloramphenicol acetyl transferase (CAT) gene³ generating a number of different constructs (see below). To study the effect of this GRE sequence without the influence of transcription factors other than those binding to the tk promoter sequence, we had to delete a fragment from all of our constructs and from ptkCAT. This fragment (see legend to Fig. 2a) contains a functional transcription factor binding site present in pUC8-, 12- or 18-derived plasmids that cooperates with the tk promoter or inserted sequences (our unpublished observation). Transient expression of the construct containing GRE but no CACCC-box sequences (pGtkCAT) in mouse L cells revealed only a very low induction upon treatment with dexamethasone (Fig. 3a).

The effect of the CACCC box on the GRE sequence was studied with the pC4GtkCAT plasmid and derivatives thereof (Fig. 2a). pC4GtkCAT contains the TO-CACCC sequence (-461 to -438) in the orientation identical to that upstream of the TO gene, separated from the MMTV-GRE (-170 to -186) by 4 bp. This construct is about sixfold inducible by dexamethasone (Fig. 2b). Furthermore, a number of CACCC/GRE spacing mutants were analysed (see legend to Fig. 2a). The level of induction shows a cyclical pattern: several peaks are apparent, corresponding to spacings of approximately 10 bp (arrows in Fig. 2b). In two of the spacing mutants showing peak inductions (pC12GtkCAT and pC37GtkCAT) the CACCC/GRE separation corresponds closely to that observed between CACCC and the two receptor binding sequences in the wildtype TO gene (Fig. 2a). Spacings of more than 70 bp showed only weak cooperativity without an obvious 10 bp periodicity (data not shown). Several control plasmids were also generated. pC29LtkCAT contains a linker of identical length instead of the GRE; pC29G*tkCAT contains one point mutation in each of the palindromic halves of the GRE (GATCAGATCAGTTTGAAACCA), a type of mutation that has been shown to abolish GRE function completely¹². pC*29GtkCAT contains the mutated TO-CACCC box, which does not compete for binding (Fig. 1). pC*(-91)29GtkCAT contains the TO-CACCC sequence with only one nucleotide changed (TCCAATACAGGCTGGGTTTGGCTCT); a corresponding mutation upstream of the β globin gene (position -91) has been shown to reduce CACCC-box dependent tran-

* Permanent address: Department of Biochemistry, Aichi Medical University, Nagakute, Aichi 480-11, Japan

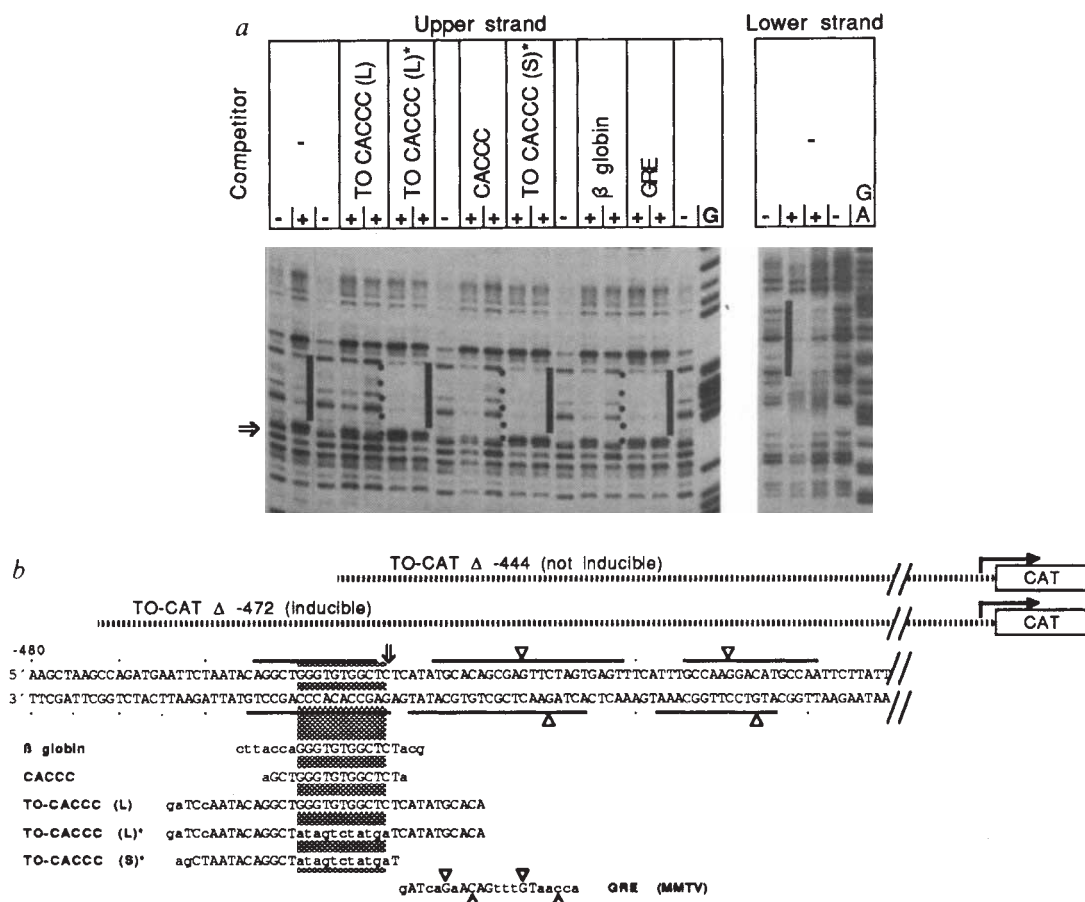


Fig. 1 Footprint experiments with different competitors. Reactions were carried out on the upper and lower strand. **a**, Footprint reactions with (+) or without (-) nuclear extract are shown and compared with the different competitors as indicated. Double stranded competitor oligonucleotides at concentrations 300-fold (left lane) and 1,600-fold (right lane) higher than the radio-labelled DNA were included in the preincubation of extract with poly(dI-dC). The A + G or the G sequencing products¹⁹ are shown as molecular weight markers. The position more sensitive to DNase I digestion upon incubation with nuclear extract is indicated by an arrow. **b**, Sequences tested for glucocorticoid inducibility. Lines above and below the CACCC box indicate the DNase footprint obtained using total nuclear extract. Also shown are the footprint regions protected by purified receptor, with close contact points as determined by DMS protection (Δ , ∇)⁷. Oligonucleotide sequences used in cloning and competition experiments are shown with those nucleotides identical to the TO sequence in capital letters and the homology to the β globin CACCC box by the shaded box. The TO deletion mutants indicating the importance of the CACCC box for steroid induction are schematically drawn.

Methods. Footprinting reactions were performed with HeLa nuclear extracts^{20,21} essentially as described²² except that the standard reaction contained 100 ng of poly(dI-dC) during the preincubation step. The DNase I digestion was carried out in the presence of 40 ng of calf thymus DNA and 1–20 ng of DNase I (Worthington) at 25 °C for 90 s, stopped by addition of 7 volumes of 400 mM sodium acetate, 30 mM EDTA, 270 ng μ l⁻¹ tRNA and frozen in liquid nitrogen. Nucleic acids were extracted with phenol-chloroform then ether, ethanol-precipitated, dissolved in 99% formamide, 10 mM EDTA with tracking dyes and loaded on a 5% polyacrylamide/7 M urea gel.

scription⁹. The deletion or mutation of the GRE abolished inducibility and deletion or mutation of the CACCC box led to only weak inductions compared with pC29GtkCAT (Fig. 3a). This proves that both elements need to be present and functional for full induction to occur. The cooperative effect seen at the functional level is not seen at the DNA binding level: footprints of two spacing mutants with low and high inducibility (pC40GtkCAT and pC29GtkCAT) are indistinguishable (Fig. 2c).

To investigate the importance of the arrangement of the different elements, and of the number of CACCC boxes, we constructed and tested recombinants which contain the CACCC box between GRE and tk promoter in either sense or antisense orientation or which contain two CACCC sequences (Fig. 3a, pG37C(a)tkCAT, pG30CtkCAT and pC29G37C(a)tkCAT). The CACCC-GRE distances were chosen to fit with one of the peaks observed in Fig. 2b. These plasmids are all equally well inducible

(Fig. 3a). In some of the constructs the distance of the GRE to the RNA start was changed (CACCC-antisense constructs). Therefore we tested directly for any effect on transcription initiation in the recombinants showing that the correct (+1) RNA start site was used, and confirming and refining the CAT assay data (Fig. 3b). These results suggest that the CACCC-box/GRE unit works independently of its orientation relative to the promoter, and that the presence of a second CACCC box does not further increase the inducibility. Since we did not test the optimal distance between the GRE and the CACCC box in the orientation used in pG30CtkCAT, we do not know whether the inducibility can be further increased by changing the position of the CACCC box.

The cooperativity of the glucocorticoid receptor and the CACCC-box binding factor suggest a protein-protein interaction leading to induced transcription. The fact that a CACCC-box footprint can be seen in the absence of a GRE footprint,

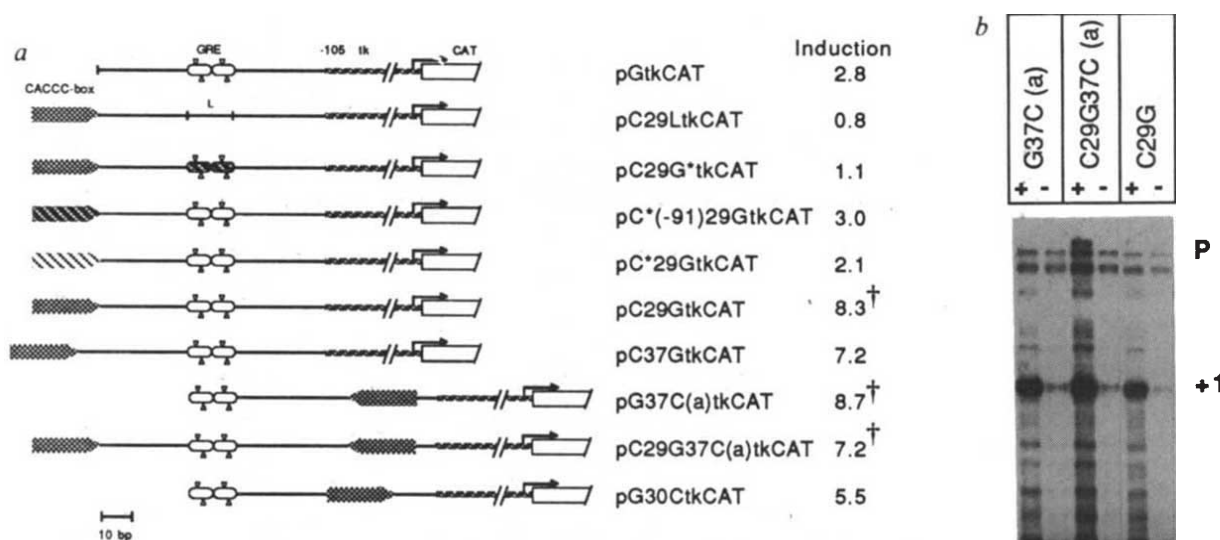


Fig. 3 Different arrangements of CACCC box and GRE. *a*, Several constructs fused to the tkCAT unit are schematically drawn to scale (bar indicates 10 bp). The CACCC box is presented by a shaded box with an arrow-head indicating 5'-3' orientation (as found in the TO upstream sequence). The MMTV-GRE sequence orientation is identical in all plasmids shown, the position being indicated by two ovals representing the TGTCT consensus sequences. Levels of induction were determined after cotransfection with pRSVGR²⁵ as described for Fig. 2b. *b*, RNase mapping of selected constructs (marked[†] in *a*). After transfection of the indicated plasmids L cells were incubated with dexamethasone (+) or left untreated (-).

Methods. pGtkCAT was generated by deleting the CACCC box of pC4GtkCAT (Fig. 2a). The plasmid pC29LtkCAT was made by removing the GRE sequence and inserting an oligonucleotide of the same length. The constructs containing a CACCC box between GRE and tk-promoter were generated by insertion of the TO-CACCC sequence (-459 to -438) in sense or antisense orientation in the *Bam*HI restriction site of pC29GtkCAT. The antisense orientation generated plasmid pC29G37C(a)tkCAT. Deletion of the distal CACCC box generated pG37C(a)tkCAT and pG30CtkCAT, respectively. For mutant CACCC or GRE sequences see text. RNase mapping was carried out with the tk-CAT probe²⁷.

functional units could lead to the strong induction observed in some systems.

We thank Dr Keith R. Yamamoto for pRSVGR, Dr Miguel Beato for the MMTV-GRE and Dr Ronald Mertz for synthesising the oligonucleotides with the technical help of Dorit Weigand. We also thank Dagmar Wolf and Karin Schulz for their excellent technical assistance, Bernd Gloss for help in the footprinting technique and Dr Michael Cross and Scott Robertson for critically reading the manuscript. This work was supported by the Bundesministerium für Forschung und Technologie and by the Deutsche Forschungsgemeinschaft.

Disruption of a putative Cys-zinc interaction eliminates the biological activity of the *Krüppel* finger protein

Norbert Redemann, Ulrike Gaul & Herbert Jäckle

Max-Planck-Institut für Entwicklungsbiologie,
Spemannstr. 35, D-74 Tübingen, FRG

Received 5 October 1987; accepted 25 January 1988.

1. Renkawitz, R., Schütz, G., von der Ahe, D. & Beato, M. *Cell* **37**, 503-510 (1984).
2. Steiner, Ch., Müller, M., Banihmad, A. & Renkawitz, R. *Nucleic Acids Res.* **15**, 4163-4178 (1987).
3. Lee, W., Haslinger, A., Karin, M. & Tijan, R. *Nature* **325**, 368-372 (1987).
4. Miksicek, R., Borgmeyer, U. & Nowock, J. *EMBO J.* **6**, 1355-1360 (1987).
5. Miksicek, R. *et al. Cell* **46**, 283-290 (1986).
6. Speck, N. A. & Baltimore, D. *Molec. cell. Biol.* **7**, 1101-1110 (1987).
7. Danesch, U., Gloss, B., Schmid, W., Schütz, G., Schüle, R. & Renkawitz, R. *EMBO J.* **6**, 625-630 (1987).
8. Dierks, P., van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J. & Weissmann, C. *Cell* **32**, 695-706 (1983).
9. Myers, R. M., Tilly, K. & Maniatis, T. *Science* **232**, 613-618 (1986).
10. Scheidereit, C., Geisse, S., Westphal, H. M. & Beato, M. *Nature* **304**, 749-752 (1983).
11. Payvar, F. *et al. Cell* **356**, 381-392 (1983).
12. Klock, G., Strähle, U. & Schütz, G. *Nature* **329**, 734-736 (1987).
13. Jantzen, H.-M. *et al. Cell* **49**, 29-38 (1987).
14. Théveny, B., Bailly, A., Rauch, C., Rauch, M., Delain, E. & Milgrom, E. *Nature* **329**, 79-81 (1987).
15. Jones, K. A., Kadonaga, J. T., Rosenfeld, P. J., Kelly, T. J. & Tijan, R. *Cell* **48**, 79-89 (1987).
16. Dorn, A., Bollekens, J., Staub, A., Benoist, C. & Mathis, D. *Cell* **50**, 863-872 (1987).
17. Chandler, V. L., Maler, B. A. & Yamamoto, K. R. *Cell* **33**, 489-499 (1983).
18. Karin, M., Haslinger, A., Holtgreve, H., Cathala, G., Slater, E. & Baxter, J. D. *Cell* **36**, 371-379 (1984).
19. Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
20. Dignam, J. D., Lebowitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475-1489 (1983).
21. Wildeman, A. G., Sassone-Corsi, P., Grundström, T., Zenke, M. & Chambon, P. *EMBO J.* **3**, 3129-3133 (1984).
22. Cereghini, S., Raymondjean, M., Carranca, A. G., Herbolme, P. & Yaniv, M. *Cell* **50**, 627-638 (1987).
23. Tabor, S. & Richardson, C. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4767-4771 (1987).
24. Wigler, M., Pellicer, A., Silverstein, S. & Axel, R. *Cell* **14**, 725-731 (1978).
25. Miesfeld, R., Godowski, P. J., Maler, B. A. & Yamamoto, K. R. *Science* **236**, 423-427 (1987).
26. Gorman, C. M., Moffat, L. F. & Howard, B. H. *Molec. cell. Biol.* **2**, 1044-1051 (1982).
27. Banihmad, A., Müller, M., Steiner, Ch. & Renkawitz, R. *EMBO J.* **6**, 2297-2303 (1987).

The best-characterized DNA-binding protein structure is the evolutionarily-conserved helix-turn-helix motif¹⁻³. Recently a second motif for DNA-binding proteins, the 'zinc finger', emerged from sequence analysis of TFIID⁴, a factor involved in the control of transcription of the *Xenopus* 5S RNA gene. The finger structure is based on pairs of Cys and His residues which are arranged around a tetrahedrally-coordinated zinc ion^{4,5}. This centre allows the folding of tandemly repeated 'finger loops'⁴ which are thought to specify the contact with target DNA. Zinc fingers have been observed in the DNA-binding protein domains of transcriptional activators in yeast⁶ and man (R. Tijan, personal communication) and in several regulatory proteins of *Drosophila*⁷⁻⁹ including proteins encoded by members of the gap class of segmentation genes^{8,9}. One of these, *Krüppel* (*Kr*), acts at the first level of the segmentation gene hierarchy¹⁰, and its protein product may bind to DNA¹¹. In addition, *Kr* is required for the development of the malpighian tubules, a posterior internal tissue that forms during later stages of embryogenesis^{12,13}. Here we show that a mutation which results in a conservative amino-acid exchange eliminates *Kr*⁺ function. The change occurs in a key position within the putative core structure of a finger, and supports the role of Cys in metal binding as proposed by Klug and coworkers⁴.

Mutations of the *Kr* gene can be ordered into an allelic series^{12,14,15}. The size of the gap in the segment pattern is quantitatively related to the functional *Kr*⁺ activity of the different *Kr*

Fig. 1 The Kr^9 mutant develops an amorphic Kr phenotype. *a*, Segmentation pattern of a *Drosophila* wild-type embryo. Thoracic (T1-3) and abdominal (A1-8) ventral denticle rows mark the anterior portion of each segment. *b*, Kr phenotype of the Kr^1 mutant which has the Kr gene deleted¹⁵. Note the absence of the thoracic and five abdominal segments which are replaced by a mirror image duplication of abdominal segment 6^{14,15}. *c*, Kr phenotype developed by the Kr^9 mutant embryos. Note that this phenotype is the lack of function phenotype (compare *b* and *c*). *d*, Malpighian tubule (arrow) in the posterior region of a wild-type embryo (numbers refer to abdominal segments; arrowhead, position of 'Filzkörper'. Denticle belts are out focus). *e*, The corresponding region (for labels and focus see *d*) of Kr^9 indicating the lack of malpighian tubules (for details see refs 12-14).

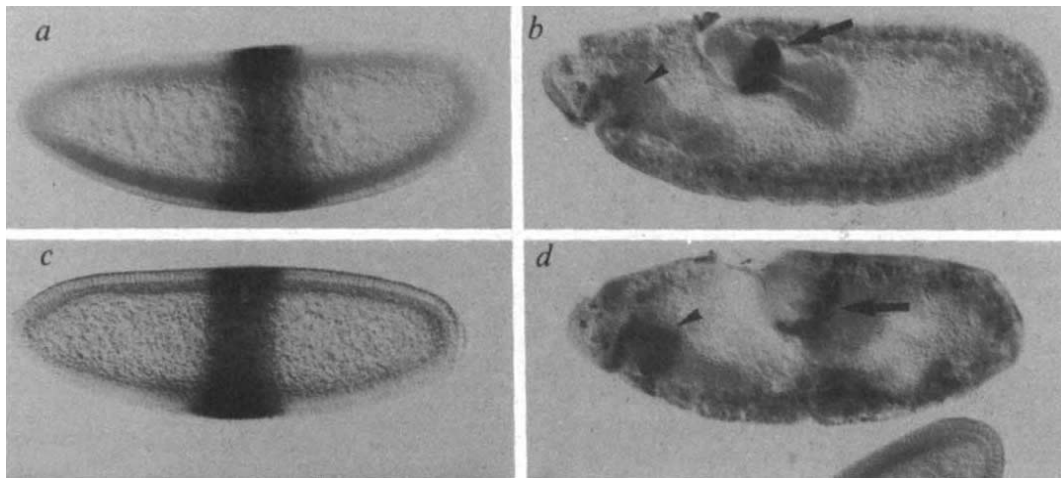
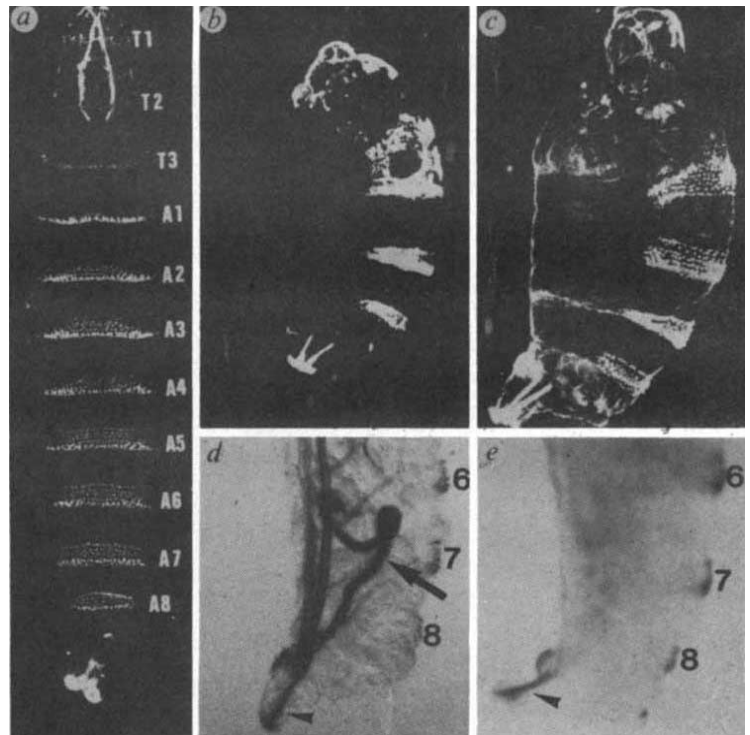


Fig. 2 Kr protein staining pattern in wild-type and Kr^9 embryos. *a*, Kr protein pattern in blastoderm stage embryos showing a band of nuclear staining in the middle region of the embryo. *b*, Kr protein staining in malpighian tubule precursor cells (arrow) during the process of outbudding during the elongated germ band stage. *c*, *d*, Kr protein staining pattern of Kr^9 embryos. Note the identical pattern at blastoderm stage (the embryo in *c* is about 20 min older than the one shown in *a*). During the elongated germ band stage (*d*), the apparent alteration of Kr protein pattern in Kr^9 embryos results because the cells which normally give rise to malpighian tubules persist in their initial position and fail to undergo outfolding (arrow). Other apparent changes are due to the different morphology of wild-type and Kr mutant embryos (see refs 12-14 for details).

Methods. Whole-mount staining of embryos involving the Kr antibodies was performed as described earlier^{16,17}. Note that Kr^9 embryos cannot be distinguished by the Kr antibody staining pattern at blastoderm stage. To ensure that the one quarter of embryos homozygous for the Kr^9 mutation were present in the egg layings examined, we counterstained a portion of the embryos of Kr^9 /SM1 parents with *fushi tarazu* antibodies (a gift of S. Carroll and M. P. Scott) which show the distortion of the *fushi tarazu* staining pattern as published for the amorphic Kr alleles²⁰.

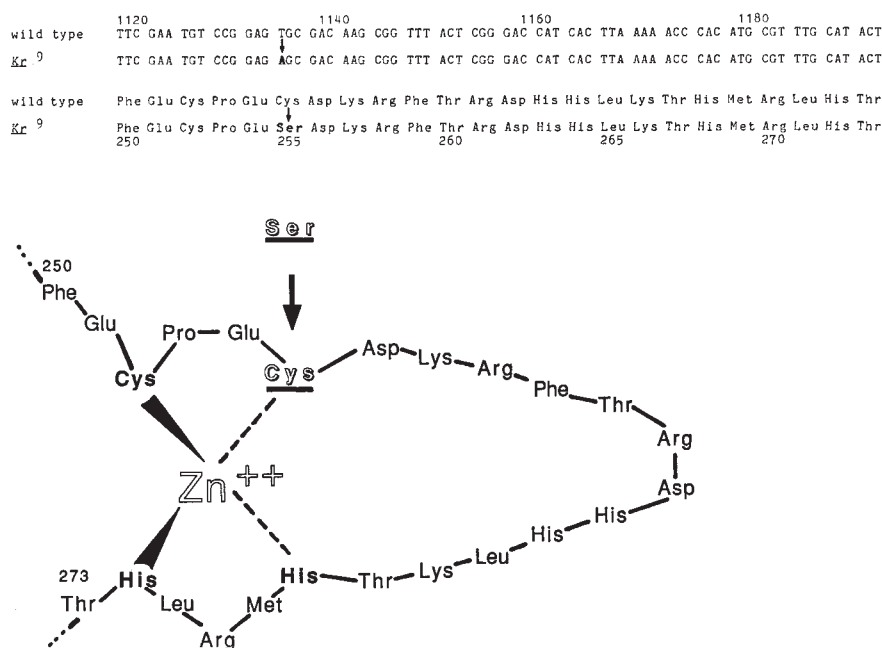
mutant embryos^{12,14,15}. In the most extreme Kr phenotype observed in Kr^- embryos ('lack of function phenotype'; Fig. 1*b*), the thoracic and five of the eight abdominal segments are deleted and are replaced by a mirror-image duplication of the sixth abdominal segment. In addition, such embryos lack the malpighian tubules in the otherwise normal posterior region¹²⁻¹⁴. To search for Kr alleles which carry their mutation in the coding

rather than in the *cis*-regulatory regions, we have examined the distribution and cellular localization of protein in Kr mutant embryos¹⁵ by use of antibodies directed against the Kr wild-type protein^{16,17}.

The Kr^9 mutation¹⁵ causes the extreme Kr phenotype (Fig. 1*c*, *e*). The distribution and nuclear localization^{11,16} of Kr protein, as visualized by the Kr antibody staining, is not changed

Fig. 3 The Kr^9 mutant is a point mutant leading to an exchange of Cys to Ser. Partial sequence of Kr wild-type and Kr^9 DNA. Note the single base-pair difference (transversion T to A; small arrow) between the Kr wild-type (first line; for the complete sequence see Fig. 2 in ref. 8) and Kr^9 sequence (second line) which results in a Ser residue in place of a Cys within the second finger region of the putative Kr finger protein domain (lines 3 and 4). The position of the exchange of the Cys to the Ser residue within the second putative Kr finger structure is schematically shown below (for orientation and the numbering of amino acids see Figs 2, 3 in ref. 8). Note that the exchange occurs in the finger core structure (possibly involving a zinc metal; not experimentally shown for Kr protein) that is required for the folding of the DNA-binding finger loop as proposed in the model by Klug and coworkers⁴.

Methods. The Kr^9 DNA was isolated as 9 kb *EcoRI* fragment from a genomic DNA library which had been prepared from the DNA obtained from Kr^9 /CyO flies¹⁵; Kr^9 DNA can be distinguished from CyO DNA by a restriction site polymorphism within the Kr gene (a *SalI* site; see ref. 15). Sequencing was carried out⁸ with two independent Kr^9 clones covering the sequences between the TATA-box and the first polyadenylation signal (Fig. 2 in ref. 8). To distinguish between trivial polymorphisms between Oregon R DNA (the Kr wild-type sequence; ref. 8) and the Kr^9 mutant DNA, we have sequenced the Kr^9 parental DNA in parallel. No difference was observed except that shown in a.



significantly in Kr^9 embryos (Fig. 2). Thus, the *cis*-regulatory elements for the spatial and temporal patterns of Kr gene expression are left unaffected by the Kr^9 mutation. This suggests that the mutation may cause a rather subtle alteration of the normal Kr protein which, however, completely eliminates Kr function as indicated by the Kr lack of function phenotype observed in Kr^9 embryos (Fig. 1).

Sequence analysis of the Kr^9 DNA revealed a single base-pair difference with respect to the wild-type sequence (Fig. 3; for the complete Kr wild-type sequence see ref. 8). This point mutation, a transversion from T to A, replaces a Cys residue by Ser (Fig. 3a). This exchange is located within the second of the four putative zinc fingers of the Kr protein (Fig. 3b). As this conservative amino-acid replacement eliminates the Kr^+ function completely, the SH group of Cys cannot be functionally replaced by the OH group of Ser. This suggests that the Ser-Cys replacement may prevent either the formation of a disulphide bridge or, in the model of finger structure proposed by Klug and coworkers⁴, the binding of a metal ion such as zinc. Zinc binding is essential for the tetrahedrally-coordinated core structure (and requires two properly positioned pairs of Cys and His residues; Fig. 3) around which a finger loop can fold. The resulting finger structure is thought to bind to about six nucleotides, half a double-helical turn of DNA^{4,18}.

Recently, a second motif for metal-binding fingers has been observed in the members of the steroid/thyroid receptor family (see ref. 19). This finger motif is based on a core structure of four Cys residues, so that the pair of His residues (Fig. 3b) is replaced by a second pair of Cys¹⁹. Our results clearly show that a replacement of Cys in the core structure of the TFIIB-related finger motif (Fig. 3) eliminates the biological activity of the protein, possibly by preventing the formation of the tetrahedrally-coordinated zinc centre. This could result in the lack of Kr function directly or it may eliminate the biological activity due to an improper folding of the entire Kr finger domain caused by the mutation of the second finger.

Note added in proof: A similar study on ADRI²¹, a positive regulator of the ADH2 gene in yeast, revealed that mutations of the Cys or His residues result in non-functional protein. The changes in ADRI, however, were less subtle than the replace-

ment of a SH-group by an OH-group, a relatively minor alteration which completely eliminates Kr function.

We thank our colleagues in the laboratory for various contributions, in particular Drs S. M. Cohen and M. P. Pankratz for critically reading the manuscript. U.G. is a Deutsche Studienstiftung fellow. The work was supported by the Deutsche Forschungsgemeinschaft (Leibniz-Programm).

Received 15 December 1987; accepted 19 January 1988.

1. Pabo, C. O. & Sauer, R. T. *Rev. Biochem.* **53**, 293-321 (1984).
2. McGinnis, W. *et al. Cell* **37**, 403-408 (1984).
3. Laughon, A. & Scott, M. P. *Nature* **310**, 25-31 (1984).
4. Miller, J., McLachlan, A. D. & Klug, A. *EMBO J.* **4**, 1609-1614 (1985).
5. Brown, R. S., Sander, C. & Argos, P. *FEBS Lett.* **186**, 271-274 (1985).
6. Hartshorne, T. A., Blumberg, H. & Young, E. T. *Nature* **320**, 283-287 (1986).
7. Vincent, A. *NAR* **14**, 4385-4391 (1986).
8. Rosenberg, U. B. *et al. Nature* **313**, 336-339 (1986).
9. Tautz, D. *et al. Nature* **327**, 383-389 (1987).
10. Akam, M. *Development* **101**, 1-22 (1987).
11. Ollio, R. & Maniatis, T. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5700-5704 (1987).
12. Jäckle, H. *et al. CSH Symp. Quant. Biol.* **50**, 465-473 (1985).
13. Gloor, H. *Arch. Julius Klaus-Stift. Vererbungsforsch.* **29**, 277-287 (1954).
14. Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Dev. Biol.* **104**, 172-186 (1984).
15. Preiss, A. *et al. Nature* **313**, 27-32 (1985).
16. Gaul, U., Seifert, E., Schuh, R. & Jäckle, H. *Cell* **50**, 639-647 (1987).
17. Gaul, U. & Jäckle, H. *Cell* **51**, 549-555 (1987).
18. Rhodes, D. & Klug, A. *Cell* **46**, 123-132 (1986).
19. Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. *Nature* **330**, 444-450 (1987).
20. Carroll, S. B. & Scott, M. P. *Cell* **45**, 113-126 (1986).
21. Blumberg, H., Eisen, A., Sledziewski, A., Bader, D. & Young, E. T. *Nature* **328**, 443-445 (1987).

Erratum

Parallel gradualistic evolution of Ordovician trilobites

Peter R. Sheldon

Nature **330**, 561-563 (1987).

In Fig. 4 legend of this letter, the symbols for the means with 95% confidence interval and number of measurements should read $\bullet_n$ and $\ominus_n$.

Timely tips for the laboratory

March is a good month for new product news: many companies launch new products in the late winter and early spring. Below are some of the best and most innovative.

FOR keeping Kimwipes, Parafilm and disposable gloves close at hand, Research Products International offers its Lab-Stak **laboratory disposables organizer** (*Reader Service No. 100*).



Multi-purpose organizer from RPI.

Constructed of sturdy acrylic, the \$99 (US) unit takes up only 11 × 7 inches on a laboratory bench. If space is a problem, the 10-pound unit can be mounted on a wall. The Lab-Stak has a top section divided into four storage compartments for loose items such as transfer pipettes, tips and micro-tubes.

Heto Lab Equipment of Denmark has a **drying chamber** for use with its line of benchtop freeze dryers (*Reader Service No. 101*). The model CD3003 is a complete unit with one removable shelf, a control box with a quick-release connection, and tubes for the compressed air supply. The polycarbonate cover lifts off easily for



Heto's freeze-dryer attachment.

access to the chamber, and the shelf is adjustable for accommodating vials of various sizes. Several trays with vials may be stacked in the dryer, as long as the total height of the stack does not exceed 115 mm. The pneumatically operated pressure plate is operated with a single button in the control box: when pressed the piston moves downwards, and when released the piston returns to its position at the top of the chamber.

Microbial tools

Sterilin offers a **four-sided inoculating loop** to facilitate dilution streaking without flaming or changing loops (*Reader Service No. 102*). Sterilin's Quadloop is a sterile

disposable polystyrene loop colour-coded according to size: 10 µl loops are blue, and 1 µl loops are white. Consisting of a loop at one end and a sphere at the other, the Quadloop provides up to four sterile surfaces by turning the handle 90° after each streaking. Sterilin says the Quadloops reduce the likelihood of cross-contamination or pathogenic aerosols. The Quadloops come 1,000 to a package, for £29.95 (UK).

Labsystems has modified its Bioscreen **microbiological growth analyser** to include an incubator that is continuously adjustable from 15 °C to 55 °C (*Reader Service No. 103*). The new model Bioscreen-C can test up to 200 samples at a time, making it useful for screening multiple samples. The £29,000 (UK) Bioscreen-C is available with software programs for routine food

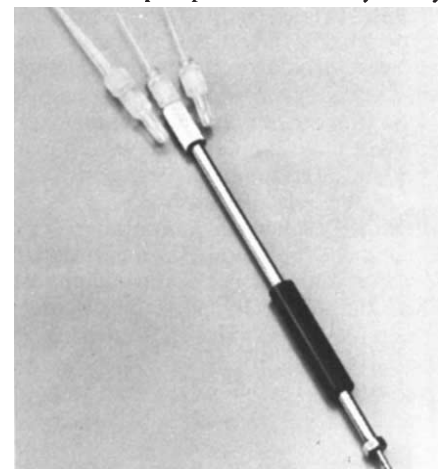


Microbiological growth analyser from Labsystems.

quality control, research on anti-microbial compounds, bioassays and food testing. Labsystems says the Bioscreen-C works as effectively with yeasts and moulds as it does with anaerobic and aerobic bacteria.

The French company Armor Equipement Scientifique has a machine that **prepares sterile media** for bacteriological cultures and fills and stacks up to 700 petri dishes in an hour (*Reader Service No. 104*). The instrument automatically mixes dehydrated media and purified water at a pre-selected temperature, deposits the media in petri dishes using a peristaltic pump, and stacks the filled dishes. The operator need only be present during the filling stage. The instrument fills the dishes under a UV lamp to ensure sterility, and because the agar is poured at temperatures lower than those necessary for manual methods, less condensation appears on the plates. Armor Equipement Scientifique offers two models: the S7000 has a 6-litre capacity, and the S12000 holds 12 litres. Both accommodate petri dishes from 50 to 150 mm in diameter as well as 120 mm square dishes.

Sterile disposable **glass capillaries for microinjection** are new from Zeiss, for use with their micromanipulation/microinjection systems (*Reader Service No. 105*). The Femtotip capillaries are very finely



Zeiss's microinjection capillaries.

pulled capillaries for the injection of DNA, RNA, dyes, toxins or other materials into tissue culture cells, and eliminate the need for researchers to pull their own capillaries. The Femtotips come with tip opening diameters of 0.5 µm and 0.2 µm. A plastic holder permits attachment of the Femtotips into the capillary holder, and an internal filament in each capillary facilitates loading. The capillaries may be filled by aspiration through the tip, immersion of the base, or pipetting with microloaders. Femtotips are available individually sealed in packs of 20 or 100, at \$186 and \$930 (US), respectively.

Gels, columns and films

The FilmRemover from Pharmacia LKB allows the **separation of thin gels** of up to 20 × 20 cm from plastic support films

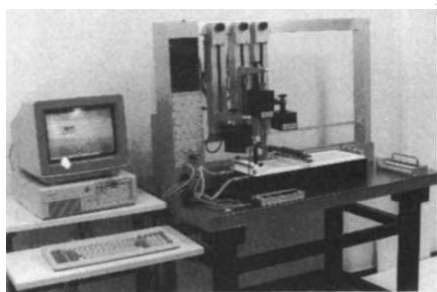


The FilmRemover from Pharmacia separates the gel from the support.

(*Reader Service No. 106*). Pharmacia says the FilmRemover allows fragile gels to be handled without damage, and facilitates their transfer to electrophoretic blotters.

A programmable temperature gradient **HPLC column heater** is new from the French firm Cluzeau Info-Labo (*Reader Service No. 107*). Cluzeau Info-Labo's column heater has a column oven with a clam shell opening that can hold HPLC columns of all lengths and head configurations. The oven has opposing metal meshed beds to guarantee a uniform temperature spread while allowing air to circulate freely. The central module of the column heater has an electronic controller that Cluzeau Info-Labo says can regulate the column temperature with a reproducibility of 0.5 °C. The gradient temperature control allows the oven to be programmed for heating in stages, with a temperature range of between subambient temperature and 99 °C, and time intervals of up to 99 hours, 59 minutes and 59 seconds.

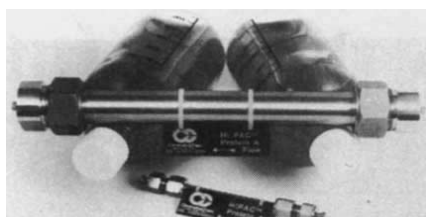
The KSV Chemical Corporation from Finland has a computerized instrument for the fabrication of **monomolecular films**, the KSV 5000 LB Center (*Reader Service No. 108*). The instrument is an



The answer for Langmuir-Blodgett thin films.

automatic Langmuir-Blodgett system for the preparation of high-quality Langmuir films and for the unsupervised deposition of ordered layers of monomolecular films. An optional alternate layer trough allows alternate multilayers to be built up in any stacking order in x, y or z form. The IBM AT-compatible system software allows the user to monitor and manipulate any of the parameters during the run, and to store data on hard disks for later processing and analysis. The surface pressure for deposition can be programmed to increase or decrease, in a linear or non-linear fashion. Deposition speeds, delays and the area of substrate to be deposited with film are also programmable.

The **HiPAC protein A columns** for affinity chromatography from ChromatoChem are designed for the purification of antibodies and other proteins (*Reader Service No. 109*). The HiPAC columns are based on a proprietary silica formulation: ChromatoChem says the columns offer the compatibility advantages of soft gels with the speed and resolution of silica HPLC, and purification times of as little as 10 minutes. The surface chemistry of the HiPAC columns involves a chemically stable linker structure that can tolerate



HiPAC protein A columns from ChromatoChem.

both acidic and basic environments without leaching. ChromatoChem says the unique chemistry also cuts down on non-specific binding and reduces column fouling. The HiPAC columns are available in analytical, bench-scale and production-scale sizes, and cost \$475–6,000 (US).

The DNA/protein connection

Promega's Gene Regulator Affinity Binding (GRAB) system takes a novel approach to the analysis and purification of **DNA binding proteins** (*Reader Service No. 110*). The system uses a *lac* repressor/ β -galactosidase fusion protein as an affinity label for recombinant plasmids containing both a *lac* operator and a cloned target sequence. The repressor moiety of the fusion protein is complexed with the plasmid and can be precipitated by the addition of a monoclonal anti- β -galactosidase antibody bound to beads. Promega says this forms an affinity matrix capable of purifying any proteins which bind specifically to the cloned target sequence, so sequence-specific DNA binding factors can be isolated from crude extracts in one step. The GRAB system sells for \$265 (US), and contains 100 μ g *lac* IZ fusion protein, 2 mg anti- β -galactosidase monoclonal antibody and 2 μ g plasmid DNA.

Invitrogen has a **cDNA library construction system** that it says can be used to construct a cDNA library of more than one million primary clones in a eukaryotic expression vector in as little as four days (*Reader Service No. 111*). The \$695 (US) Librarian I system uses Invitrogen's Megaclone nonpalindromic linker insertion technology, and a pCDM8 vector with a CMV eukaryotic promoter/enhancer, a T7 promoter, a polyadenylation signal and splice segment, SV40 and polyoma virus origins, an STVS (ColE1-



Libraries of up to one million clones are possible with the Librarian I system from Invitrogen.

like) high copy plasmid origin, and an M13 origin. Invitrogen says the Megaclone technology increases insertion efficiency and lowers background to less than one per cent, without methylation, cutback or dephosphorylation of the vector.

The Dutch company Pierce says its **Micro BCA★ protein assay reagent** offers sensitive measurement of protein solutions from 0.5 μ g ml⁻¹ to 20 μ g ml⁻¹ (*Reader Service No. 112*). The Micro BCA★ reagent is based on bicinchoninic acid, and is meant to be used with a spectrophotometer. Pierce says the reagent is effective in measuring low concentrations of protein, even in the presence of ionic and non-ionic detergents. Pierce's reagent is a pre-formulated aqueous system that is supplied with three separate components. The final working reagent is stable for 24 hours when all three components are mixed in the appropriate ratios.

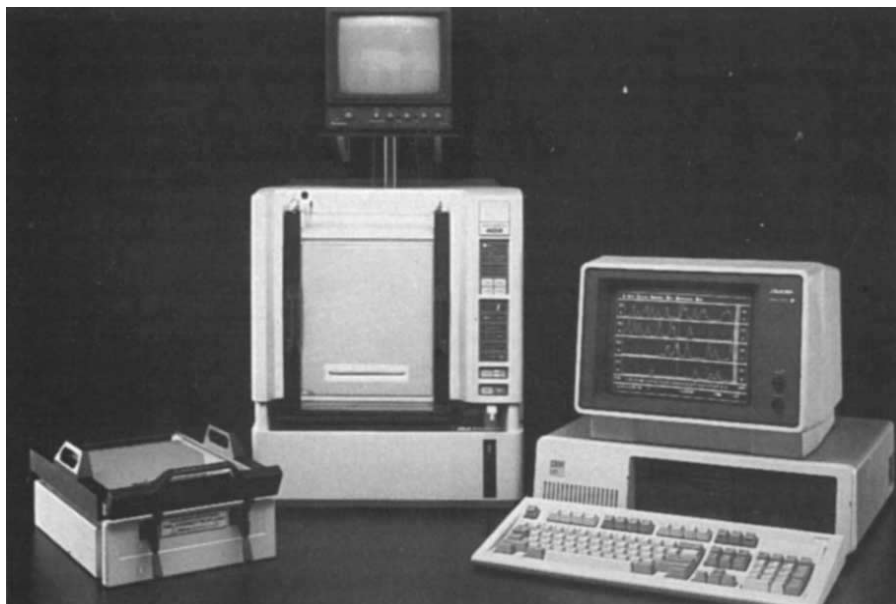
The **polymerase chain reaction** technique for amplifying specific DNA sequences is now available as a complete system from Perkin-Elmer Cetus (*Reader Service No. 113*). Perkin-Elmer Cetus' \$8,500 (US) DNA Amplification System



PCR becomes automated with the Perkin-Elmer Cetus system.

consists of two GeneAmp DNA amplification reagent kits and a DNA thermal cycler. Using the system, researchers can amplify, *in vitro*, a target DNA sequence by over 100,000-fold within hours. The GeneAmp kit includes thermostable polymerase from *Thermus aquaticus*, dNTPs, optimized buffers and a control template with matching primers to verify kit performance. The microprocessor-controlled DNA thermal cycler holds 48 0.5-ml microfuge tubes, and can store up to 93 protocols for temperature profiles.

Inhibit-Ace is the name of a new **RNase inhibitor** from 5 Prime → 3 Prime, Inc. that it says is 50 times more potent than human placental RNase inhibitor in



Automated DNA sequencing using radionucleotides from EG & G.

cDNA synthesis, *in vitro* translations and *in vitro* transcriptions (*Reader Service No. 114*). 5 Prime → 3 Prime, Inc. says that one unit of Inhibit-Ace has been shown to inhibit 90 per cent of the activity of up to 250 ng of pancreatic RNase A. The Inhibit-Ace RNase inhibitor costs \$50 (US) for 250 units.

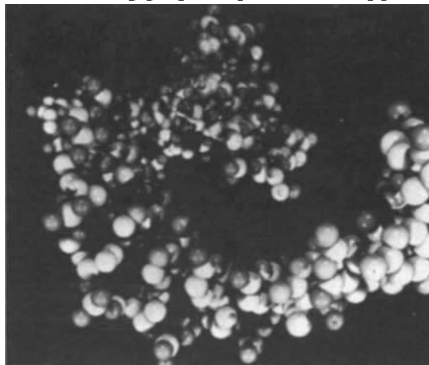
The latest automated DNA sequencer on the scene is the Acugen 402 system from EG & G Biomolecular (*Reader Service No. 115*). The \$45,000 (US) Acugen 402 uses traditional radioisotopic labelling techniques and eliminates the need for gel autoradiography by directly determining ³²P-labelled nucleic acid fragments from the polyacrylamide gel. Raw sequence data is displayed on a video monitor during the run, and then is transferred to a personal computer for analysis and base determination. The system's gel casting device does away with clamps and tape. The Acugen 402's IBM PC-compatible software automatically assigns bases, corrects for gel aberrations and provides special notation for sequencing abnormalities so they can be given closer scrutiny. EG & G says the Acugen 402 can read approximately 250 bases per clone and six clones per gel run. Each run takes less than seven hours, so two runs can be achieved per day.

Computing

Environmental toxicity information can now be accessed online, through the QSAR system addition to the Numerica network from Technical Database Services (*Reader Service No. 116*). The QSAR system has three main options: the "properties" function contains information on the physical or chemical properties of organic compounds, the "ecotox" function performs quantitative assessments of the fates and biological effects of chemi-

cals in the environment, and the "acquire" function allows access to the US Environmental Protection Agency's compilation of 96,000 aquatic toxicity tests for 3,900 chemicals and 2,500 species. The only input required for identifying a chemical is its CAS registry number or molecular structure. The system can be used for screening substances for potential biological activity, and for the evaluation of new chemicals and pesticides.

Ardent Computer Corp. will soon be releasing Dynamic Object-Rendering Environment (Doré) high-level graphics software for use with its Titan supercomputer, the Cray supercomputers and the Sun workstations (*Reader Service No. 117*). The Doré software library lets users describe a scene, produce complex images from the scene data, and manipulate the massive amounts of data interactively and dynamically. Using the Doré hierarchical scene data base, abstract scenes can be created using graphics primitives, appear-



Ardent's software does molecular modelling.

ance attributes, geometric attributes, studio objects, and organizational objects that can be collected in a data base and called back into a scene. Applications for the Doré software include computational physics, computational fluid dynamics,

seismic data processing, reservoir simulation, imaging and animation. Ardent will be licensing the software to universities and research laboratories for \$250 (US), and to commercial users for \$15,000 (US) plus support fees. □

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.

ADVERTISEMENTS

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

POB 126, D-8000 Munich 65, FRG

Reader Service No.49



HIGH QUALITY AGAR and AGAROSE

FUNAKOSHI PHARMACEUTICAL CO. offer you two kinds of High Quality AGAR for culture media, and five kinds of High Quality (high gel strength, high porosity, exact electroendosmosis and high purity) AGAROSE for electrophoresis, immunoelectrophoresis, immunodiffusion, isoelectric focusing, etc. with reasonable packages.



For detailed information please contact:

FUNAKOSHI PHARMACEUTICAL CO., LTD.

2-3 Surugadai, Kanda, Chiyoda-ku, Tokyo, Japan

Telephone : Tokyo 03-293-2367

Telefax : 81-3-295-5545

Telex : J28489 FUNA

Reader Service No.53

Changing prospects for UK job market

Richard Pearson

Britain's Chancellor Nigel Lawson, some say, is presiding over an economic 'boom'. But what likelihood is there that the recovery will result in employment opportunities?

MARCH is the month of the UK budget when the Chancellor of the Exchequer reports on the financial state of the country and makes his major changes to taxes. This year there is a prospect of major tax cuts because of booming tax revenues as the economy continues its expansion for the fifth successive year. Although this is a long recovery period by historical standards, industrial output has only just returned to the level of 1979 while unemployment, at 2.7 million under the current counting system, is still exceptionally high. The figure becomes worse still when account is taken of the half million or more on special schemes and the fact that the new method of counting deflates the total relative to the figures given in the 1970s. So how does the boom in the economy translate itself into jobs?

The early 1980s saw a massive shakeout from manufacturing, with more than 2 million jobs lost in the decade to 1985, an annual loss rate of 3 per cent. Much of that contraction was structural as businesses closed and uneconomic capacity was withdrawn. As a result, there was a significant growth in UK productivity as the activities with the lowest productivity have been removed. During the period to 1995, this job loss is expected to continue, albeit at a slower rate, as companies seek further improvements in competitiveness by improving working methods, investing in new technology and continuing to cut out uneconomic capacity. So although

Employment change 1986-95 (thousands)

Primary and utilities	-114
Manufacturing	-467
Construction	+169
Distribution, transport and communications	+337
Business and other services	+1,369
Public services	+115
Total	+1,406

output is expected to rise, productivity will go up more rapidly and jobs will continue to be lost. The long-term decline in employment in agriculture and the utilities is expected to continue, reflecting longer-term productivity gains.

Construction, by contrast, is likely to continue its recovery and create jobs, reversing the losses of the past decade. The main growth sectors will continue to be in retailing, hotels and catering, and particularly financial and business services, which will generate more than a million new jobs by 1995, an annual growth rate of

3 per cent. In the public sector, modest growth is expected, driven by health and welfare needs. Self-employment will also keep on growing into the 1990s benefiting, as will the growing number of small companies, from further subcontracting by the larger companies and the public sector. Across the whole economy, an extra 1.4 million jobs are expected.

Unfortunately for the unemployed, the working population will also continue to grow until 1990 before it starts to decline. The most dramatic component of change will be the massive decline in the number of workers aged under 24, down by a quarter, which will be good for youth unemployment but which will also play havoc with recruiters seeking school-leavers for jobs, training and places in further and higher education (see *Nature* 329, 470; 1987). Women's participation in the labour market will also continue to increase, partly because of more part-time job opportunities. The net effect will be a fall in unemployment of about 600,000 over the decade to 1995, when the number of registered claimants will total about 2.5 million. Women will continue to be the beneficiaries of job growth, much of which will be in part-time jobs; their unemployment rate in 1995 will be about 6 per cent compared with 11 per cent for men. The figures for women will, however, continue to be depressed by the many women who would take work if available but do not register for benefits.

At an occupational level, job growth will continue to benefit managerial and professional workers while operative and labouring jobs will continue to disappear. About 0.9 million professional jobs will be generated, an annual increase of 1.75 per cent. It is estimated that by 1995 almost one in four of the labour force will be engaged in such jobs, and they will be the largest occupational grouping. The next largest will be sales and personal services, accounting for one in five, followed by managers, craft, clerical and operatives (see figure). The growth in the number of managers will be caused largely by structural changes whereby the growth in services generates a large number of small units, each headed by a manager, although the proportion of managers in business and related services will also rise. By contrast, the growth in professional occupations will result from their increasing their share of employment in all sectors, most notably in health, education

and public administration where they will account for about half of all the jobs.

How then does this translate into messages for the job-seeker looking for new qualifications? The first point is that there is no universal linkage between occupations and qualifications. Second, the United Kingdom, like most other countries, has no systematic data on the skill content of jobs, let alone how the skills needed are changing. The evidence on qualifications shows many examples where there is no formal link with occupa-



Occupational distribution predicted for 1995. Source for figure, Review of the Economy and Employment, Institute for Employment Research, 1987.

tions. For example, whereas nurses and doctors have to have relevant qualifications, only a minority of accountants have accountancy degrees. In information technology, electronics engineers need relevant degrees, usually in electronics or physics, but only a minority of programmers and systems analysts have computer science degrees. Moving on to the vexed question of what makes a manager, where masters' degrees in business administration are rare, there is not even a common language for the skills involved, job descriptions drawing on such indefinable attributes as leadership, analytical skills, problem-solving and communicative skills. There are many thousands of qualifications in the United Kingdom, some leading to specific careers, others of a more general nature. Very few relate to each other. A map is needed of existing qualifications and their comparability and value, after which an attempt could be made to link them to the changing occupational opportunities and skill needs of the economy. Until then, the job-seeker trying to adjust to the new opportunities will be floundering when it comes to choosing a relevant qualification. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Brighton BN1 9RF, UK.